

Fig. 1
1994

RESULT 5
PCPAL3
LOCUS
DEFINITION
P. crispum mRNA for phenylalanine ammonia-lyase 3.
X81159
ACCESSION
X81159.1 GI:535007
VERSION
X81159.1
KEYWORDS
deaminase; PAL-3 gene; phenylalanine ammonia-lyase.
SOURCE
ORGANISM
Petroselinum crispum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
1 (bases 1 to 2403)
Apert, C., Logemann, E., Hahlbrock, K., Schmid, J. and Amthelp, N.
Structural and catalytic properties of the four phenylalanine
ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.)
Eur. J. Biochem. 225 (1), 491-499 (1994)
JOURNAL
MEDLINE
95010141
PUBMED
7925471
REFERENCE
2 (bases 1 to 2403)
Apert, C.
Direct Submission
Submitted (22-AUG-1994) C. Apert, Institut fuer
Pflanzenwissenschaften, Eidgenossische Technische Hochschule,
Zuerich, Universitaetsstrasse 2, 8092 Zurich, SWITZERLAND
JOURNAL
TITLE
AUTHORS
FEATURES
Source
1. 2403
/organism="Petroselinum crispum"
/db_xref="taxon:4043"
89. 2245
/gene="PAL3"
89. 2245
/EC_number="4.3.1.5"
/note="tetramere subunit"
/product="phenylalanine ammonia-lyase 3"
/codon_start=1
/protein_id="CA57057.1"
/db_xref="GI:535008"
/db_xref="SWISS-PROT:P45729"
/translation="MAYVNGITNGHANGNGLDLCKMKEDPLNMGVAAEAATGSHLDEV
KRMVAEYKRPVVKLEGETLITISOVAISARDGSGVVELSEERAGVAKASDVMDSM
NKGDSYGVTFGATSHRRTKGGALQKELIRFLNAGIFGSGAEGNNTPHSATRA
ANLVIRINILLOGYSIGIREILEITKELFNHNTPCPLRGITIASGDLVPLSYIAGLL
TGRPSKAVGPTGVTLSPEFAFKLAGVGGFELOPKREGIALVNGTAVGSGMASVYLE
ENITLAVLAEVMSAIFAEVMQCKPEPTDHLTKLKHHPGOIEAAIIMEHLLDGSAYVK
AAOKLHENDPLOKPKODRYALRTSPOMLGPOIETVIRSTKMIEREINSVNDPLIDVS
RNKAIHGNGFQSPIGVSMDNPLAIAIGKLMFAQFSELVNDPYNGLDPSNLSCGN
PSLDYGFKAELIAIASYCELOFLANPVTNHOVSAEQHODVNSLGLISSRTSEAVE
IKIMSTPELVGLCOAIDLRLHEENIKSTVKNVSOVAKRYLTWGVNGEELHPSRPECV
DLRYVDREYITPAIDPCSATYPLMOKLRETLVEHALNNGDKERNLSTSIPOKIAAF
EDBLKALLPKREYETARALESQNPAPNRIKEDRSYPLKRVREELCTEYLTGCKVVS
PCDEFKVFYTAASKGELIDPLECLSNNGAPLPIC"

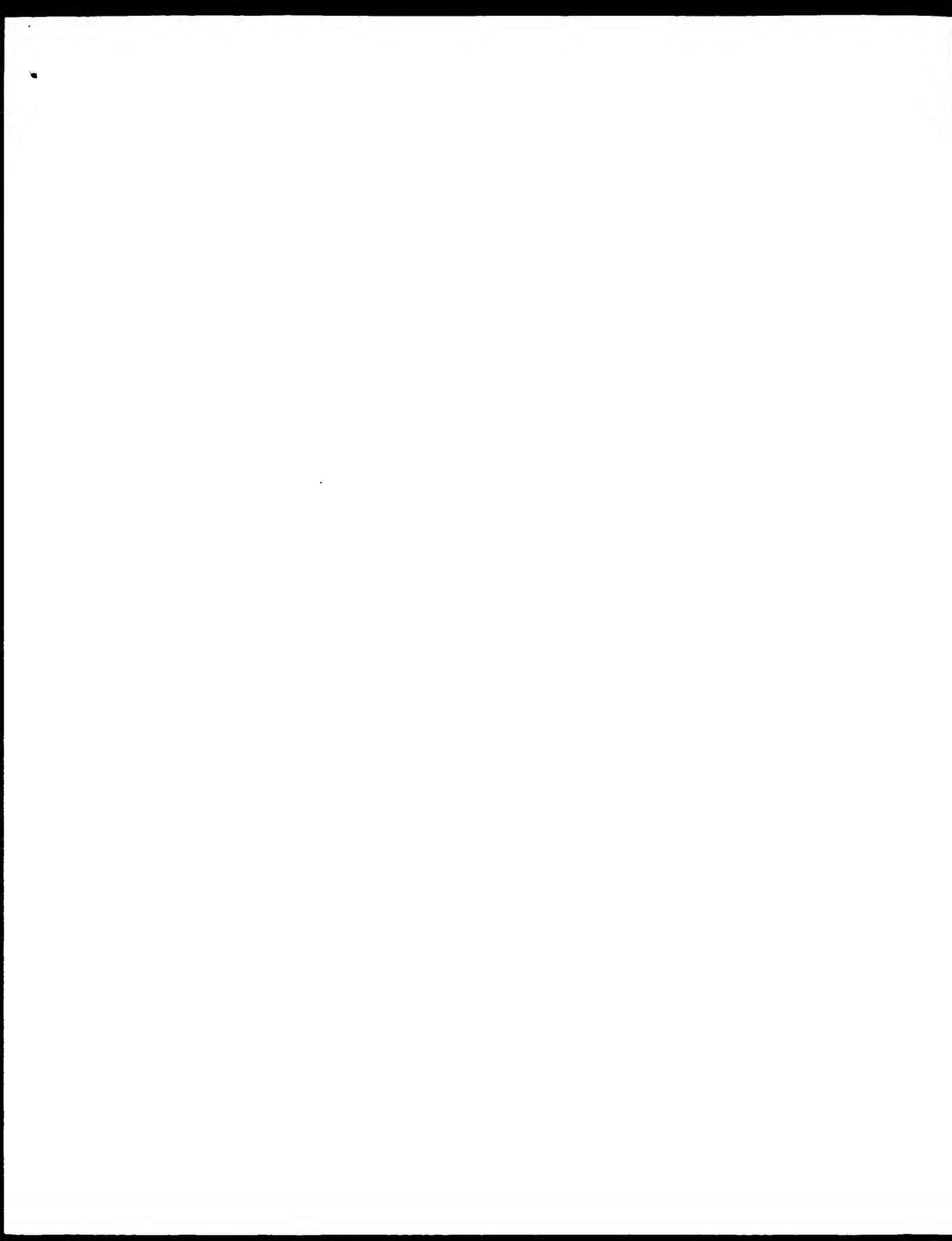
BASE COUNT 708 a 461 c 573 g 661 t
ORIGIN

Query Match 52.0%; Score 1270.8; DB 8; Length 2403;
Best Local Similarity 74.2%; Pred. No. 11e-293;
Matches 1634; Conservative 0; Mismatches 562; Indels 6; Gaps 2;



135 ACCTTAATGAGTCTGTTAATGAGTTCATGATCAAGATCCATTGAACTGGGAGTTGCAAG 194
126 ACAGGAAAGGATTGATTTGTGATGATGAAAGAAAGATCTTTGAACTGGGAGGAGCTG 185
195 CGAGAGCGTTGACCGGAAAGTCACTTGATGAGGTGAAGAAAGATGGTGGGAGTTGCAAA 254
186 CGAGAGCATTTGACAGGAGATCATTTGACAGAAATTAAGAGATGGTGGTGAATGACAGA 245
255 AGCGGCTGTGAAGTCCGAG 314
246 AGCGGCTGTGAAGTCCGAG 305
315 CTGCTAATGACATGACACCGTGAAGGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 374
306 CTGAGGATGATAGTGTGAG 362
375 AGCGGATGATGAGTGTGAG 434
363 AGCGGATGATGAGTGTGAG 422
435 CCAGCGGCTGTGAG 494
423 CTAGTGGCTGTGAG 482
495 AGCTCAATGATTTTGAAG 551
483 AGCTTATGAGTCTTGAAG 542
552 CACTTCACATTCAG 611
543 CATTACCACTCCGAG 602
612 GTTACCGCGGATCGGATTCGAGATTTGAGAAAGATCCAGAGTTCCTTGAAGAGAGAG 671
603 GCTATTCAGAGATCGGATTCGAGATTTGAGAAAGATCCAGAGTTCCTTGAAGAGAGAG 662
672 TCACCGCTGTGAG 731
663 TTAAGCTGTGAG 722
732 CATACATCGCGGCTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 791
723 CCAACATCGCTGAG 782
792 AGCTCAATGATTTTGAAG 851
783 TAAAGCTCAG 842
852 TACAGCGGAG 911
843 TACAGCGGAG 902
912 CCATGCTGTGAG 971
903 CTATGCTGTGAG 962
972 TCAGTGTGATGAG 1031
963 TCAGTGTGATGAG 1022
1032 ATACAGCTGTGAG 1091
1023 ACATATCGAG 1082
1092 ACCTCAAG 1151
1083 ACCTTAAAG 1142
1152 GTATGCTGTGAG 1211
1143 GATATGCTGTGAG 1202

1212 CAAGCAAAATGATCGAG 1271
1203 CAAGCAAAATGATCGAG 1262
1272 CAAGCAAAATGATCGAG 1331
1263 CAAGCAAAATGATCGAG 1322
1332 CAAGCAAAATGATCGAG 1391
1323 CAAGCAAAATGATCGAG 1382
1392 CAAGCAAAATGATCGAG 1451
1383 CAAGCAAAATGATCGAG 1442
1452 CAAGCAAAATGATCGAG 1511
1443 CAAGCAAAATGATCGAG 1502
1512 CAAGCAAAATGATCGAG 1571
1503 CAAGCAAAATGATCGAG 1562
1572 CAAGCAAAATGATCGAG 1631
1563 CAAGCAAAATGATCGAG 1622
1632 CAAGCAAAATGATCGAG 1691
1623 CAAGCAAAATGATCGAG 1682
1692 CAAGCAAAATGATCGAG 1751
1683 CAAGCAAAATGATCGAG 1742
1752 CAAGCAAAATGATCGAG 1811
1743 CAAGCAAAATGATCGAG 1802
1812 CAAGCAAAATGATCGAG 1871
1803 CAAGCAAAATGATCGAG 1862
1872 CAAGCAAAATGATCGAG 1931
1863 CAAGCAAAATGATCGAG 1922
1932 CAAGCAAAATGATCGAG 1991
1923 CAAGCAAAATGATCGAG 1982
1992 CAAGCAAAATGATCGAG 2051
1983 CAAGCAAAATGATCGAG 2042
2052 CAAGCAAAATGATCGAG 2111
2043 CAAGCAAAATGATCGAG 2102
2112 CAAGCAAAATGATCGAG 2171
2103 CAAGCAAAATGATCGAG 2162
2172 CAAGCAAAATGATCGAG 2231
2163 CAAGCAAAATGATCGAG 2222
2232 CAAGCAAAATGATCGAG 2291
2223 CAAGCAAAATGATCGAG 2282
2292 CAAGCAAAATGATCGAG 2351
2283 CAAGCAAAATGATCGAG 2342



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 19:53:15 ; Search time 6234 Seconds

(without alignments)
11400 233 Million cell updates/sec

Title: US-09-964-992a-3

Perfect score: 2442
Sequence: 1 ggcgaatcgtatcaatcccc.....aaaaaaaaaaaaaaaaaaaaa 2442

Scoring table: IDENTITY_NTC
Gapop 10.0 ; Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GenBank

1: gb-ba:
2: gb-ba:
3: gb-in:
4: gb-cm:
5: gb-cv:
6: gb-pat:
7: gb-ph:
8: gb-pl:
9: gb-pr:
10: gb-ro:
11: gb-sl:
12: gb-sy:
13: gb-un:
14: gb-vi:
15: em-ba:
16: em-fun:
17: em-hum:
18: em-in:
19: em-mu:
20: em-om:
21: em-or:
22: em-ov:
23: em-pat:
24: em-ph:
25: em-pl:
26: em-ro:
27: em-sl:
28: em-un:
29: em-vi:
30: em-hq-hum:
31: em-hq-inv:
32: em-hq-other:
33: em-hq-mus:
34: em-hq-pla:
35: em-hq-rod:
36: em-hq-mam:
37: em-hq-vrt:
38: em-sy:
39: em-hqo-hum:
40: em-hqo-mam:
41: em-hqo-other:

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	2442	100.0	2442	8	AF299330 Lactuca s
2	1550.2	63.5	2322	8	Y12461 Helianthus
3	1286.2	52.7	2444	8	AF401636 Rehmannia
4	1280.6	52.4	2375	8	X17462 P. cristum R
5	1270.8	52.0	2403	8	PCPAL4
6	1259.2	52.0	2448	8	AF346115
7	1260.8	51.6	2462	8	NTPEPAL
8	1254.2	51.4	2330	8	CHEPAL
9	1250.8	51.2	2527	8	TOBPAL1
10	1247.4	51.1	2495	8	AB042520
11	1235.2	50.6	2428	8	SSNPAL
12	1234	50.6	2403	8	DLJ002221
13	1231.4	50.4	2396	8	IRBPALA
14	1227.8	50.3	2384	8	PCPAL2
15	1227.4	50.3	2441	8	DB3075
16	1220.4	50.0	2352	8	DB3076
17	1219.6	49.9	2380	8	POPPALGA
18	1219	49.9	2414	8	AF092957
19	1217.2	49.8	2439	8	AF337955
20	1215.8	49.8	2406	8	AF367308
21	1210.2	49.6	2527	8	AF045919
22	1208.8	49.5	2154	8	AV134695
23	1206.2	49.4	2335	8	IPHPAL
24	1204.6	49.3	2309	8	AY079363
25	1203.2	49.3	2178	6	AX412245
26	1203.2	49.3	2178	6	AX412319
27	1203.2	49.3	2178	6	AX412320
28	1202.8	49.3	2409	8	MSPAL
29	1201.6	49.2	2253	8	AY036011
30	1200.6	49.2	2402	8	AF088956
31	1199.2	49.1	2448	8	AF036948
32	1194.6	48.8	2607	8	CL043338
33	1178	48.2	2427	8	CCL38753
34	1173.2	48.0	2471	8	CCL38754
35	1170.4	47.9	2367	8	PEAPAL
36	1159	47.5	2607	6	EA04042
37	1158.6	47.4	2276	8	AF250845
38	1150.2	47.1	2501	8	AF337954
39	1102.6	45.2	2536	8	AY096657
40	1093.8	44.8	2249	8	BPPAL
41	1089.4	44.6	5465	8	DB5850
42	1081.8	44.3	2044	8	PA016130
43	1080	44.2	R0917	8	AF004502
44	1067.6	43.7	4472	8	PEAPAL2
45	1061.4	43.5	6306	8	AB008200

ALIGNMENTS

RESULT 1
AF299330 2442 bp mRNA linear P1N 31-DEC-2001
DEFINITION Lactuca sativa phenylalanine ammonia-lyase mRNA, complete cds.
ACCESSION AF299330
VERSION AF299330.1 GI:16001006
KEYWORDS
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
Asterales; easterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
REFERENCE
1 (bases 1 to 2442)
Campos-Vargas,R., Nonogaki,H., Suslow,T. and Saltveit,M.

TITLE Characterization of phenylalanine ammonia-lyase (PAL) gene in
JOURNAL wounded lettuce leaf tissue
REFERENCE 2 (bases 1 to 2442)
AUTHORS Campos-Vargas, R., Monogaki, H., Suslow, T. and Saltveit, M.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-2000) Vegetable Crops, University of California
Davis, One Shields Ave., Davis, CA 95616, USA

FEATURES
source 1..2442
location/Qualifiers
/organism="Lactuca sativa"
/db_xref="taxon:4236"
119..2254
/note="PAL1"
/codon_start=1
/product="phenylalanine ammonia-lyase"
/protein_id="AAU55242.1"
/db_xref="GI:18001007"
/translation="MENGHNHNGVNVNELIKDPLNMGVAEAITGSHLDEVKMAAEF
RKPVKLSGETLTVSVAGIAANDSTIVELSEARVGVKMSDMVESNKGIDS
YGVITGATSHRTKOGALOKELIRLNAGLPGQTESHTLPHSATRAAMIYAIN
TLLOGSGIRPELLEATTKPLNNITPCLPLRGTITASGLVPLSYIACILGRPSK
AVSGEVLNAEKAFKAFAGVGGPEFLQPEGLALVNGTAVGSGMASVLPANVATL
LSEVLSHIFAEVWGKPEFLDHLTHKIHIPQIEAALIKETILDSQYKAAQVHE
MDLQKPKDRTALRISPVLDPLIEVIRKSTIKIEREINSDVNDPLIDVSRNALHG
GNFGTPIGVSMOTRLAIAIGKLMFQPSSELYNNGLEPNSISGRNPSLDYGF
KGGEIAMAQSELOFLANPVTHVOSAFQHNQVNSLIGIISARKTAFLVDILKMS
TYVALCOSIDLRLHENMKSTVKNVTSQAVKDELNGVLEPKEDILRVYD
REYFAVLDVQSGVTPLMOKLROVLVDHALNNGELENKNTNLSIPKIPFEEELKVL
LPKEVEVRITAVENDTISTPRIRKATRSYLYLFVVEBELRPSLTREKVTSPGEEDP
VFIMACKQIILDPLECLGGMNEPDLIC"

BASE COUNT 693 a 553 c 592 g 604 t
ORIGIN

Query Match 100.0%; Score 2442; DB 8; Length 2442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAATCTGATCAATACCATTCACGACAAAGAGTGTGAGTCTAGTGTGTGAAGAGT 60
DB 1 GAGCAATCTGATCAATACCATTCACGACAAAGAGTGTGAGTCTAGTGTGTGAAGAGT 60
QY 61 ACACAAATTGATTTGTTCTTTGTTTATGATGATGATGATGATGATGATGATGATGAT 120
DB 61 ACACAAATTGATTTGTTCTTTGTTTATGATGATGATGATGATGATGATGATGATGAT 120
QY 121 GGAGAAACGCTAATCACTTAATGAGTGTGATGATGATGATGATGATGATGATGATGAT 180
DB 121 GGAGAAACGCTAATCACTTAATGAGTGTGATGATGATGATGATGATGATGATGATGAT 180
QY 181 CTGGGAGCTTGCAGCGGAGCGCTTACCGGAAGTCACCTTATGAGTGAAGAAGATGCT 240
DB 181 CTGGGAGCTTGCAGCGGAGCGCTTACCGGAAGTCACCTTATGAGTGAAGAAGATGCT 240
QY 241 TGGGAGCTCAGAAAGCGGCTGTGAAGCTCGAGAGAGAGAGCGCTTACAGTTTTCAGGT 300
DB 241 TGGGAGCTCAGAAAGCGGCTGTGAAGCTCGAGAGAGAGAGCGCTTACAGTTTTCAGGT 300
QY 301 GCGGGGATGCGAGCTGTCTAATGACAGTGAACCGTGAAGTGAAGTGTGCGAAGCGCG 360
DB 301 GCGGGGATGCGAGCTGTCTAATGACAGTGAACCGTGAAGTGAAGTGTGCGAAGCGCG 360
QY 361 GAGGCTGAGATTAAAGCGGAGTGTGAGTGTGATGAGAGCATGAAATTAAGGAAGTGA 420
DB 361 GAGGCTGAGATTAAAGCGGAGTGTGAGTGTGATGAGAGCATGAAATTAAGGAAGTGA 420
QY 421 TAGTTATGCTGACACCGGCTTGCAGCACTCTACGAGAAATTAAGCAAGCGG 480
DB 421 TAGTTATGCTGACACCGGCTTGCAGCACTCTACGAGAAATTAAGCAAGCGG 480
QY 481 TCGTTTACGAGAGAGCTCATTAGATTTTGAACCGCGGAATATTGCGCAATGGAACGGA 540
DB 481 TCGTTTACGAGAGAGCTCATTAGATTTTGAACCGCGGAATATTGCGCAATGGAACGGA 540

QY 541 AACAGGACACACTTGCATTCAGCAGGAGAGCGGCGATGATGATGATGATGATGATGATGAT 600
DB 541 AACAGGACACACTTGCATTCAGCAGGAGAGCGGCGATGATGATGATGATGATGATGATGAT 600
QY 601 CCGTCTGCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 CCGTCTGCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 720
DB 661 TAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 720
QY 721 TGTCCATTAT 780
DB 721 TGTCCATTAT 780
QY 781 CCGTCTGCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CCGTCTGCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 GTTCTTCAGATTACAGCGGAGAGAGGCTTACACTTGTTAACGGGACCGGCTGGGCTC 900
DB 841 GTTCTTCAGATTACAGCGGAGAGAGGCTTACACTTGTTAACGGGACCGGCTGGGCTC 900
QY 901 CCGGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 CCGGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 ATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 ATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 CAAATTTGAAGCATCACCCTGCTCAAAATCGACCGCGCGCATGATGATGATGATGATGATGAT 1080
DB 1021 CAAATTTGAAGCATCACCCTGCTCAAAATCGACCGCGCGCATGATGATGATGATGATGATGAT 1080
QY 1081 CGGAAGCATTTACGTCAGAGCGGCGCAAAAGGCTCCAGAAATGAGACCGCTTACAGAAAC 1140
DB 1081 CGGAAGCATTTACGTCAGAGCGGCGCAAAAGGCTCCAGAAATGAGACCGCTTACAGAAAC 1140
QY 1141 AAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 AAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 AATCGATTCATCAACCAAAATGATGAGAGGGAATCAATTCGTCACGACCAACCATTT 1260
DB 1201 AATCGATTCATCAACCAAAATGATGAGAGGGAATCAATTCGTCACGACCAACCATTT 1260
QY 1261 GATGAGCTTTCGAAACAAAGCTTTTACAGTGTGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 GATGAGCTTTCGAAACAAAGCTTTTACAGTGTGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 AGTTTCATGAGCAACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 AGTTTCATGAGCAACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 ATTTTCGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 ATTTTCGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 ACGTAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 ACGTAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TGTGAGCTTGCAGTTCGCAAAATCGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1560
DB 1501 TGTGAGCTTGCAGTTCGCAAAATCGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1560
QY 1561 CAATCAAGACGTTAATTTCTGCGATTAATTTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 1561 CAATCAAGACGTTAATTTCTGCGATTAATTTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1620

[illegible]

QY 441 GCTTGGGCGCCACCTCTTCCAGGAGAACTAAAGCAAGCGCGTCTTTACAGAGAGGCA 500
 Db 441 GATTGGTCTCTCTCTCTAGAGAGAGTAAAGATGGTGGTCTTACAAAAGAGCTCA 500
 QY 501 TTAGATTTTGAACCGCGGAAATATTCGCAATGAGCAAGCAAGCAAGCAAGCTTCAC 560
 Db 501 TTAGATTTCTGAACCGCGGAAATATTCGCAAGCAAGCAAGCAAGCAAGCTTCAC 560
 QY 561 ATTGAGCCACAGAGCGGATGATGTCAGATTAACACCTCTCTCAGAGGTACTCG 620
 Db 561 ATTGAGCCACAGAGCGGATGATGTCAGATTAACACCTCTCTCAGAGGTACTCG 620
 QY 621 GCATCGATTTGAGATCTTGAAGCATACCAAGTCTCTTAAACAACTACACCTT 680
 Db 621 GCATCGATTTGAGATCTTGAAGCATACCAAGTCTCTTAAACAACTACACCTT 680
 QY 681 GTTTACCCCTCCGTCGAAACCATCAGCCCTCCGCTGACCTTGTCCATTAATCATG 740
 Db 681 GTTTACCCCTCCGTCGAAACCATCAGCCCTCCGCTGACCTTGTCCATTAATCATG 740
 QY 741 CCGGCTCTTAACGCGCGCGCCCACTCAAAAGCGGTGGCGCCACGAGAAAGTCTCA 800
 Db 741 CCGGCTCTCTAACGCGCGCGCCCACTCAAAAGCGGTGGCGCCACGAGAAAGTCTCA 800
 QY 801 ATGCCGAAAGGCGCTTCGTCGAGCGGAGTTGAAGTGGTCTTCCAGTTACAGCGCA 860
 Db 801 ATGCCGAAAGGCGCTTCGTCGAGCGGAGTTGAAGTGGTCTTCCAGTTACAGCGCA 860
 QY 861 AAGAGGGCTAGCAGCTTGTAAACGCGACCGCTGGGCTCCGGATGGCTTCCATGTTTC 920
 Db 861 AAGAGGGCTAGCAGCTTGTAAACGCGACCGCTGGGCTCCGGATGGCTTCCATGTTTC 920
 QY 921 TATTGATCTATGACTTGGTGTGTGGAGTGTAAACGCGCTTCCAGTTACAGCGCA 980
 Db 921 TATTGATCTATGACTTGGTGTGTGGAGTGTAAACGCGCTTCCAGTTACAGCGCA 980
 QY 981 TTATGCAAGGAAAGCGCGATTTACCGATCACTTACACAAATGAGCATCACTG 1040
 Db 981 TTATGCAAGGAAAGCGCGATTTACCGATCACTTACACAAATGAGCATCACTG 1040
 QY 1041 GTCGAATGAGAGCGCGGATGATGAGTATTTTGGAGAGCAAGCAAGTACCTCAAG 1100
 Db 1041 GTCGAATGAGAGCGCGGATGATGAGTATTTTGGAGAGCAAGTACCTCAAG 1100
 QY 1101 CCGCGCAAAAGGTCCAGCAAAATGAGACCGGTTACAGAAACCAAAACAGATGTTATGTC 1160
 Db 1101 CCGCGCAAAAGGTCCAGCAAAATGAGACCGGTTACAGAAACCAAAACAGATGTTATGTC 1160
 QY 1161 TCCGTATCATCTGCCATGAGTGGCTGAGCTGCAATGCAATGCAATGCAATGCAATG 1220
 Db 1161 TCCGTATCATCTGCCATGAGTGGCTGAGCTGCAATGCAATGCAATGCAATGCAATG 1220
 QY 1221 TGAATGAGAGGAGATCAACTCTCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1280
 Db 1221 TGAATGAGAGGAGATCAACTCTCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1280
 QY 1281 AAGCTTTACAGCGGTGTAATTCAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1340
 Db 1281 AAGCTTTACAGCGGTGTAATTCAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1340
 QY 1341 GTTCGCCATTTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1400
 Db 1341 GTTCGCCATTTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1400
 QY 1401 ATTCTACAAACATGATTAACATCAATCTCTCGGAGAGCAAGCAAGCAAGCAAGCA 1460
 Db 1401 ATTCTACAAACATGATTAACATCAATCTCTCGGAGAGCAAGCAAGCAAGCAAGCA 1460
 QY 1461 ACGGTTCAAGAGGTGAGAAATGCAATGCTTCTTACGTGAGCTTCAAGTTCTCG 1520
 Db 1461 ACGGTTCAAGAGGTGAGAAATGCAATGCTTCTTACGTGAGCTTCAAGTTCTCG 1520

QY 1521 CAATTCAGTCCAGCAACCATGTTCAAGCGCGCAACAAACAGATCAAGAGCTT 1580
 Db 1521 CAATTCAGTCCAGCAACCATGTTCAAGCGCGCAACAAACAGATCAAGAGCTT 1580
 QY 1581 TCGAATTAATTTAGCGAGGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1640
 Db 1581 TCGAATTAATTTAGCGAGGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1640
 QY 1641 CGACATCTATGCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 1700
 Db 1641 CGACATCTATGCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 1700
 QY 1701 AATGCAAGTGAAGAAACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1760
 Db 1701 AATGCAAGTGAAGAAACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1760
 QY 1761 ACGGAGCTCCAGCGCTGAGATTTCTCGGAAAGATTTCTTCTGTTGTTGTTGTTG 1820
 Db 1761 ACGGAGCTCCAGCGCTGAGATTTCTCGGAAAGATTTCTTCTGTTGTTGTTGTTG 1820
 QY 1821 AATGCTCTGCTTACATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1880
 Db 1821 AATGCTCTGCTTACATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1880
 QY 1881 TCCGACAGGTTCTGCTGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1940
 Db 1881 TCCGACAGGTTCTGCTGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1940
 QY 1941 CCTCATCTTCTTCAAAAGATCTGCTTCTGCAAGCAAGCAAGCAAGCAAGCAAG 2000
 Db 1941 CCTCATCTTCTTCAAAAGATCTGCTTCTGCAAGCAAGCAAGCAAGCAAGCAAG 2000
 QY 2001 AAGTTGAAGGTGTGTAATGCTTATGCAATGCAATGCAATGCAATGCAATGCAAT 2060
 Db 2001 AAGTTGAAGGTGTGTAATGCTTATGCAATGCAATGCAATGCAATGCAATGCAAT 2060
 QY 2061 AAGCTTGCAGATGCAAGCGGTGTAATGCTTATGCAATGCAATGCAATGCAATG 2120
 Db 2061 AAGCTTGCAGATGCAAGCGGTGTAATGCTTATGCAATGCAATGCAATGCAATG 2120
 QY 2121 TGAACGCAAAAGGTGAGCGTGGCGGAGAGGAGTTCGACAGGCTTTCACGCGCAT 2180
 Db 2121 TGAACGCAAAAGGTGAGCGTGGCGGAGAGGAGTTCGACAGGCTTTCACGCGCAT 2180
 QY 2181 GCAAGGTCATTAATTTGATCGGTGTTGAGTGTCTTGAAGGAGTGAAGTGAAGT 2240
 Db 2181 GCAAGGTCATTAATTTGATCGGTGTTGAGTGTCTTGAAGGAGTGAAGTGAAGT 2240
 QY 2241 TTCCATATATGTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2300
 Db 2241 TTCCATATATGTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2300
 QY 2301 TTTTGTGTTT 2309
 Db 2301 TTTTGTGTTT 2308

RESULT 3
 AF401636 2444 bp mRNA linear, PLN 08-AUG-2001
 LOCUS AF401636
 DEFINITION Rehmannia glutinosa phenylalanine ammonia-lyase (PAL1) mRNA,
 complete cds.
 ACCESSION AF401636
 VERSION AF401636.1 GI:15100058
 KEYWORDS
 SOURCE Rehmannia glutinosa.
 ORGANISM Rehmannia glutinosa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Lamiales; Lamiales incertae sedis;
 Rehmannia.
 REFERENCE 1 (bases 1 to 2444)
 Yun,S.-J., Lee,B.-K., Park,M.-R., Srinivas,B. and Chun,J.-C.

OY	1674	ATTATACCCCATTTTGGAGGAAACATGAAAATGCAGCTTCGAACAACCGCTTAAGCAAGTG	1733
Db	1660	ATTTAGAGCAGCTTGGAGGAAATTBAGACTTCAGATTAAAAAACCCTTAGCCAAAGTGG	1719
OY	1734	CGAAAAAGTGTCTACCATAGGCGGTCAACGGGAGCTCCACCCGTGAGATTGTGCCAGA	1793
Db	1720	CGAAGAGCACTGTGCAATAAGCTATTAAAGGGAACTTCAACCGCTTAAGATTGCCAGA	1779
OY	1794	AAGATCTCCCTCGGTGTGTGATCTGTAAATAGTTTGTGGTTATATGAGAGCTTTGCA	1853
Db	1780	AGCATCTTCTCCCTGTGTGTGAGACCCGCAATACGCTTTTGCAATACATCACAGATCCGTGCA	1839
OY	1854	GCGGCGACATACCGATTAAATGAGAGCTTCGACAGCTTCTGTGACACGACGCTTAACA	1913
Db	1840	GCGGAGCACTACCGCTTATGATTAATAATTAGAGTAATCTGTGTAAACAGAGCTTAATA	1899
OY	1914	ACGGCGAAGCGGAGAACACTTAACACCTTCATCTTCCAAACATGCTTACCTTCGAAC	1973
Db	1900	ACGGTGCAGCTGAGAAAAAGCTGACACGCTCATTTTTCAAACATCCAGCGCTTTGACG	1959
OY	1974	AAGAAITGAAAAGTCTGTATTACGAAAAGAAATGGAAGTGTGATGAAATGCTTATAGATG	2033
Db	1960	TACAGTTGAAGCGCATCTTGTCCATAAAGAGCTGACAGTGCACGAGCGGCTTGGACATG	2019
OY	2034	ATACATTGTGCATTTCCAACACAGATTAAAGCTGTGACATCTGATACCCTGTGTATAGTTTG	2093
Db	2020	GAAATCCGCGCATTTGCTTACAGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2079
OY	2094	TAAACGAGAGAGCTCCGCGACAGAGGCTTTTTTGACDCCGAGAAAGCTACBCNCCGAGAGG	2153
Db	2080	TCAAGAGAGCAACTTGGAGAGCAACTACTTGACCGCGCAAAAGCTGCTTTTCGCGGCGAGG	2139
OY	2154	AGTTCACAGAGTGTTCATAGAGATATGTAAGAGTCAAAATATTATGATCTGTGTAGAT	2213
Db	2140	AAATGTGTAAGGTGTTCACAGCTTTGACCAAGAGTGTGATGTGTGATGATCTGTGTGAAGT	2199
OY	2214	GCTCTTGAGAGGTGCAATGCGGAAACCTCTTCCCAATATGTTAGCAAGAGAGCTGCAAAACGG	2273
Db	2200	GCTTGTAGGGTTGCAATGAGTGACCTCCGCCCTATCTGTACTTCATTTMAATTTGTGTTT	2259
OY	2274	TTTGCAATTGATTGTTGAATATCTCTGTTTTTTTTTTTTTTTTTTTTTTAAATTATTTGCAATT	2333
Db	2260	GTTGCTTAAGGACTTCTGCTTTCTTAAATGTTTTTCCCTCAAGTGTGCTTTAAATTAAATAGT	2319
OY	2334	AATATTCATCAACAAAGACTTCACACTTCAGAGTGTGATGTATGTTGTAATATATATATA	2393
Db	2320	TGATTTCTGTCTAGCGTGAATTAATTGTGTAAAAAAAAAAATATGACGTTTATGTTTTTTAA	2379
OY	2394	TTTACTTATTATTCTTCTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2442
Db	2380	AA	2428

RESULT 4

LOCUS LCPAL4

DEFINITION P. crispum RNA for PAL4, phenylalanine ammonia-lyase.

ACCESSION X17462

VERSION X17462.1 GI:20453

SOURCE pal4 gene.

ORGANISM Petroselinum crispum.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.

REFERENCE 1 (bases 1 to 2375)

AUTHORS Schulz,W., Eiden,H.G., and Hahlbrock,K.

TITLE Expression in Escherichia coli of catalytically active phenylalanine ammonia-lyase from parsley

JOURNAL FEBS Lett. 258 (2), 335-338 (1989)

MEDLINE 90092501

PUBMED 2589222

[illegible]

135 ACCTTAATGAGTCGTTAATGAGTTGTGATCAAGATGCATTCGAGTGGAGTTGCGAG 194
126 AGGGAGAGGATTAATTTGTGATGAAGAAAGAGATCTTTAACTGGGGAGTGGCTG 185
195 CGAGAGCGTTGACCGGAAGTCACTTGTAGAGTGAAGAGATGTTGCGGAGTTTCGAA 254
186 CGAGGCAATTCGACGAGTCAATTTGGACGAAGTTAAAGAGATGTGGCTGATACGAA 245
255 AGCGGTGTGAAAGTGCAGAGAGAGAGCTTACAGTTCTCTAGGTGAGGAGATCCAG 314
246 ACCCGGTGTGAGTGCAGAGAGAGAACTGACAAATTTTCAGGTGAGGAGTTCGCG 305
315 CTGCTAATGATGATACACCGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 374
306 CTAGGATATATAGT---GTGTTAAGTGAAGTTCGAGAGAGAGAGAGAGAGAGAG 362
375 AGCGGATAGTATGTTGCTTATGAGAGAGATGATTAAGAGAGTGAATAGTTATGCTCA 434
363 AGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
435 CCACCGGTGTGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
423 CTACTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
495 AGCTGATTAATTTTGAAGCGCGGATATTCGCAATGAGAGAGAGAGAGAGAGAGAG 551
483 AGCTTATTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
552 CAGTTCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 611
543 CATTACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
612 GTTATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
603 GCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
672 TCACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 731
663 TTAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 724
732 CATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
723 CCTACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
792 AAGTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 851
783 TACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
852 TACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
843 TACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
912 CCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 971
903 CTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
972 TCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
963 TCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
1032 ATACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
1023 ACCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
1092 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
1083 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
1152 GTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1211
1143 GATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1202

1212 CAACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1271
1203 CAACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1262
1272 CCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1331
1263 CCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322
1332 ACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1391
1323 ACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1382
1392 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1451
1383 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442
1452 GTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1511
1443 GCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1502
1512 AGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1571
1503 AGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1562
1572 TTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1631
1563 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
1632 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1691
1623 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1682
1692 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1751
1683 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1742
1752 TGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1811
1743 TGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1802
1812 TGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1871
1803 TGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1862
1872 TGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1931
1863 TGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1922
1932 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1991
1923 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1982
1992 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2051
1983 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2042
2052 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2111
2043 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2102
2112 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2171
2103 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2162
2172 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2231
2163 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2222
2232 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2291
2223 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2282
2292 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2333

[illegible][illegible]

Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Asteridae: Ericales: Theaceae: Camellia.

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 2330)
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (22-JAN-1994) Satoru Matsumoto, Natl. Res. Inst. of Vegetables, Ornamental Plants and Tea; Kanaya 2769, Kanaya, Shizuoka 428, Japan (tel:0547-45-4101, fax:0547-46-2169)
Location/Qualifiers
1. 2330
/organism="Camellia sinensis"
/cultivar="Yadukita"
/db_xref="taxon:4442"
/clone="A21"
/issue_type="Young leaf"
/clone_lib="lambda gtl1"
132. 2276
/codon_start=1
/product="phenylalanine ammonia-lyase"
/protein_id="BA05643.1"
/db_xref="gi:662271"
/translation="MDSTTAINGVSGSGSPGCLKPLNMGVAEAMKSHLEEVKGMVEERKRVNRLGGETLISQVLAIVRGEVAVELSESREKVSMMVMSNMKG TDSYGTTCGATGATSHRRKTKGKLOKELIPLNAGLPGNTESECHLIPQATPAALY RINTLQSTGSTRPELEIAISKLNNTTCLPLRGTITTSGLDVLSTACILTCRH NSKAVPTGELIHPKEARLAVGEGFPEIOPBGLAVGTAVSGLSMVLPEANI LAVSEVLSAIFAEVMOGKPEFTDHLHKHKGTEAAMIEHILDSVYKAAK LHEMDLPQKDRYALRSPQWGLPIEIRSTKSEIEINSDNPLINSRKA LHGFNOSTPIGMDNTPLAVASIGKLMFAPSELYNDFYNNGLSNLGGNPSLD YGFGAELIAMAAYCSSEIOLPLAVPTNHVQSAHODVNSLGISSKTAFAVDLKL MSTRYVALCOAVDRHPRENLHNTYKSTYSQAKRYLIMCVNGELHPSKTCCKDLK VVDREYIPATYIDPCSATYPLMOKLROVLVHAKNGSKRKNLSTIPKIRAFBEI KTLPEVESTRALIEGNSAIPRIIECHSRYLYKFEVREELTELLTEKVSPEE FDKVFTALCKGEMIDPLMDCLEKEMNGAPLIC"

polyA_signal
polyA_site 2330
BASE COUNT 630 a 496 c 595 g 609 t
ORIGIN

Query Match 51.4%; Score 1254.2; DB 8; Length 2330.
Best local similarity 74.7%; Pred. No. 1e-289;
Matches 1590; Conservative 0; Mismatches 533; Indels 6; Gaps 1,

QY 158 TTGTGCATCAAGCATTCATGCAAGTGGGAGCTTCGACGCGGAGCGCTTCACCCGCAAGTCAC 217
DB 186 TTTTGTCTAAAGACCTTTGATTTGAGAGTGGGAGCGGCGGAGCAATGAAGGAGATCAT 245
QY 218 CTTCATGAGTCAAGCAAGATGCTTCGAGATTCAGAAAGCCGCTGCTGAAGCTGGAGCA 277
DB 246 TTGAGAGAGTGAACGCCATGCTGGAGCATTTCCGAAAGCCAGTGGTGAAGCGG 305
QY 278 GAGAGCTTAAAGATTTCTAGTGGGAGGATGAGTGAATGATGATGATGATGATGATG 337
DB 306 GAGACTTTGACGATATTCGAGCTGGCGGAGTCCGCGG-----TGAGCGTACTGAGCTG 359
QY 338 AAGGTGAGCTGTCGGAAGCCCGGAGGCTGAGTTAAGGCGAGTACGATGATGGGTTATG 397
DB 360 GCGGTGAGCTGTCGAGTCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 419
QY 398 GAGAGCATGAATTAAGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 457
DB 420 GAGAGTGAATTAAGGAGGAGCATGATGATGATGATGATGATGATGATGATGATGATG 479
QY 458 CATGCGAGCATTAAGCAAGGAGGCTTTAAGAAAGAGCTCATTAATTTTGAAGCGC 517
DB 480 CATAGAGGAGCAAGGAGGAGGCTTCCTCAAAAGGAGCTTAATAGCTTCGAAAGCT 539

QY 518 GGAATATTGGGCAATGAGAGGAAACAAAGGCAACATCTTCATATTCATACCAAGAGCC 577
DB 540 GGAATATTGGGCAATGAGAGCAAGAGCTTCATATTCATATTCATACCAAGAGCA 599
QY 578 GCCATGATCGTCAGAAATCAACACCTCTCCAGGCTTACTCCGAGCATTCGATGAGAT 637
DB 600 GCTATGCTGTGAAGATCAACACCTCTCCAGGATATTCGAGCATTCGATGAGAT 659
QY 638 TTGGAAGCATCAACCACTTCCTTAACCAACATTCACCTTCCTTACCCCTCGGGA 697
DB 660 TTGGAAGCATTCCTAAATTCCTCAACCAACATTCACCTTCCTTACCCCTCGGGA 719
QY 698 ACCATCACCGCTCCGCTGACCTTCCTCATATCAATACCGCGCTTCATACCGG 757
DB 720 ACTATCACCGCTTCGCTGACCTTCCTCATATCAATACCGCGCTTCATACCGG 779
QY 758 CCCCCCACTCAAAAGCGTTGCGCCCAACGAGAGAGTCTCAATGAGGAGGCTTC 817
DB 780 CCGCAATTCGAAAGCGCTGCGCTTACCTGAGAGAAATTCCTCAACCGGAGGAGGCTTC 839
QY 818 GCTGAGCGGAGTGAAGGTGGGCTTCCTGAGTTACAGCGGAGAGGAGGAGGAGCT 877
DB 840 GCTGAGCTGAGAGTGAAGGTGGGCTTCCTGAGTTACAGCGGAGAGGAGGAGGAGCT 899
QY 878 GTTAACGAGCGCGCGCTGCGGCTCGGAGTGGCTTCATGTTCTATTTGATGCTAATGA 937
DB 900 GTCAACGAGCGAGCTGTTGCTGCGGCTTCGCTTCATGTTCTTCGAGGCTTACATTA 959
QY 938 CTGGGCTGTTGCGGAGTGTATCGGCGATCTTCGAGTGTATGCAAGGAGGAGGCG 997
DB 960 CTGGGCTGTTATCGAAGTTTGTACCGGATTTTCGTTGAGTTATGCAAGGAGGAGGCC 1019
QY 998 GAGTTTACGATCACTTGACACACAAATTTGAAGATCAACCTGCTCAATCGAGGCGG 1057
DB 1020 GAGTTTACGATCACTTGACACAAATTTGAAGATCAACCTGCTCAATCGAGGCGG 1079
QY 1058 GCGATCATGAGATATTTTGGAGGAGGAGGATTCAGTCAAGGCGGCGGCAAGCTCCAC 1117
DB 1080 GCGATCATGAGATATTTTGGAGGAGGAGGATTCAGTCAAGGAGGAGGAGGAGGAG 1139
QY 1118 GAAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1177
DB 1140 GAAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1199
QY 1178 TGGCTGGAGCTCAATCGAAGTAAATCGATCAACCAATGATGAGAGGAGGAGGAG 1237
DB 1200 TGGCTGGAGCTCAATCGAAGTAAATCGATCAACCAATGATGAGAGGAGGAGGAG 1259
QY 1238 AATTCGTCACGACCAACCCCATTCGAGCTTCGAGGAGGAGGAGGAGGAGGAGGAG 1297
DB 1260 AATTCGTCACGACCAACCCCATTCGAGCTTCGAGGAGGAGGAGGAGGAGGAGGAG 1319
QY 1298 AATTCGTCACGACCAACCCCATTCGAGCTTCGAGGAGGAGGAGGAGGAGGAGGAG 1357
DB 1320 AATTCGTCACGACCAACCCCATTCGAGCTTCGAGGAGGAGGAGGAGGAGGAGGAG 1379
QY 1358 ATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1417
DB 1380 ATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1439
QY 1418 TTACATGCAATCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1477
DB 1440 TTACATGCAATCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1499
QY 1478 GAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1537
DB 1500 GAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1559
QY 1538 CATGTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1597
DB 1560 CATGTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1619
QY 1598 ACGAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1657

[illegible]

Db 1116 GCTCTGAGAACATCTCCCACTGCTGCGCCCTCAGATGAGTGAATACCAAGCAAC 1175
 QY 1217 AAAATGATCGAGAGGAAATCAATTCGTCACACCAACCATTTGATGACGTTTCAGA 1276
 Db 1216 AAAATGATGAGAGGAAATCAATTCGTCACACCAACCATTTGATGACGTTTCAGA 1235
 QY 1277 AACAAAGCTTTACAGGAGTAACTTACCTCAAGCAACCCCAATTCGATTCGATGACAC 1336
 Db 1236 AACAAAGCTTTACAGGAGTAACTTACCTCAAGCAACCCCAATTCGATTCGATGACAC 1295
 QY 1337 ACCGCTCTGCGCATTTGTCATGAGAAATCAATTTGATGACGTTTCAGA 1396
 Db 1296 TCTACATTTAGGCTTCTGCTGCTGATGAGAAATCAATTTGATGACGTTTCAGA 1355
 QY 1397 AACGATTTCTACAAATGATGATTCATTCGATTCGATTCGATTCGATTCGATTCG 1456
 Db 1356 AACGATTTCTACAAATGATGATTCATTCGATTCGATTCGATTCGATTCGATTCG 1415
 QY 1457 GACTAGCGGCTTCAAGGAGTGAATGCGCATGCTTCTGATGACGTTTCAGA 1516
 Db 1416 GATTTGATTTCAAGGAGTGAATGCGCATGCTTCTGATGACGTTTCAGA 1475
 QY 1517 CTGCGAAATCGAGTCAACCAATGTCGATTCGATTCGATTCGATTCGATTCGATTCG 1576
 Db 1476 TTGGCTAATCCGCTGACCAATGCTGACGAGCGCGGAGCAACCAAGATGTTAAC 1535
 QY 1577 TCTCTGAGTAAATTCAGAGAGGAAACATTCAGATGATTCGATTCGATTCGATTCG 1636
 Db 1536 TCCCTGGGTTGATCTCTGCAAGAAAGCATGATGATTCGATTCGATTCGATTCGATTCG 1595
 QY 1637 TCGTCGACATCTAGTCTGCTCTGATGCAATTCGATTCGATTCGATTCGATTCGATTCG 1696
 Db 1596 TCCCTGACATTTAGTCTGCTCTGATGCAATTCGATTCGATTCGATTCGATTCGATTCG 1655
 QY 1697 ATGAATCGAGTGAAGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1756
 Db 1656 TTGAAGAGAGGAGTGAAGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1715
 QY 1757 GTCAAGAGGAGTGAAGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1816
 Db 1716 GTCAAGAGGAGTGAAGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1775
 QY 1817 GTGATAGCTGCTGCTGATGATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1876
 Db 1776 GCGGATAGCTGCTGATGATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1835
 QY 1877 AAGTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1936
 Db 1836 AAGTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1895
 QY 1937 AAGTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1996
 Db 1896 AAGTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1955
 QY 1997 AAGTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2056
 Db 1956 AAGTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2015
 QY 2057 ATTAAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2116
 Db 2016 ATTAAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2075
 QY 2117 TTTTGGAG 2176
 Db 2076 TTTTGGAG 2135
 QY 2177 ATGTCGAAG 2236
 Db 2136 ATGTCGAAG 2195
 QY 2237 CCTCTGCAATATCTTACAG 2296
 Db 2196 CCTCTGCAATATCTTACAG 2255

QY 2297 TGTTTT 2303
 Db 2256 GGTGTT 2262

RESULT 14
 PCPAL2
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

PCPAL2
 P. crispum mRNA for phenylalanine ammonia-lyase 2.
 X81138
 X81138.1 GI:534892
 deaminase; PAL-2 gene; phenylalanine ammonia-lyase.
 Petroselinum crispum
 Petroselinum crispum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
 1 (bases 1 to 2384)
 Appert, C., Logemann, E., Hahlbrock, K., Schmid, J., and Amrhein, N.
 Structural and catalytic properties of the four phenylalanine
 ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.)
 Eur. J. Biochem. 225 (1), 491-499 (1994)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

FEATURES
 source
 1. .2384
 /organism="Petroselinum crispum"
 /db_xref="taxon:4043"
 /note="tetramere subunit"
 59. .2209
 /gene="PAL2"
 59. .2209
 /gene="PAL2"
 /EC_number="4.3.1.5"
 /note="deaminase subunit"
 /codon_start=1
 /product="phenylalanine ammonia-lyase 2"
 /protein_id="CAA57056.1"
 /db_xref="GI:534893"
 /translation="MENGATITNGVNCNMGDMQKTEDPLVWYIAEAMTGSHTDE
 VKRMVAEYRKPVKLGCEFLITISOVAALISARGSCVYVLEISPAARAGYKASIDVWMS
 MKGTDSTGYTTFQCATSHRTRKOGALOKELIRPLNAGIGRNGSDNLTPLSATRAAM
 LVRINTLDGYSIGIREILEATITKFLNOMITPDLKGTITNSGDLVPLSTIAGLLTG
 RPSKAVGPTGVLISPEEAFKLAGVEGFEFLQPEKGLAVNGTAVGSMVLFEEA
 NIILAVLAEWASALFEAVMVGKPEFDHLHKLHHPGQIEAAAIHEHILDSAVYKKA
 OKLHEMDPLQPKRODYLALRTSPQWIGPQIEVIRSTKMIEREINSVNDPLIDVSRN
 KAIHGNFOGTPIGMSMDNTRILAIKGLMAQPSKLVNPNYNGCLPSNLGGKNS
 LIDCFKGAETIAMSVCSELOPLANTVNTVOCARQNDVNSICLISRKSPAYELL
 KLMSTPLVGLCAULDLRLHLENUKSTYANNTSYAKKYLTMVNGELHRSRFEKDL
 LRFVDEYITFATYDPSATYPLDMKLTROTLEIAIMKMDNERNNNTSITFKIATFED
 ELKALLPKEVESRAALAEKGNPAIPNR:EC:SVPL:YFVKREL:IEVLTGSEKVTSPG
 ERFDDKVIAMSKGELLIDPLECLIESNGPLDPC"

gene
 CDS

BASE COUNT 700 a 427 c 581 g 676 t
 ORIGIN

Query Match 50.38; Score 1227.8; DB 8; Length 2384;
 Best Local Similarity 72.18; Pred. No. 228; 283;
 Matches 1661; Conservative 0; Mismatches 622; Indels 20; Gaps 4;

QY 135 ACGTTAATGAGCGGTGAATGAGTTCGATCAAGCATTCATTCGAGCGGAGTTCGAG 194
 Db 99 ACGGAACGAGGATGATTTGATGAAGACCGAAGATCCGTGTGATCGAGGAATCGCGG 158
 QY 195 CGAGCGCGTTCACCGAAGTACCTGATGAGTGAAGAGATGTTGCGGAGTTTCAGAA 254
 111111 1111 11111 11111 11111111111111 1111 11

QY 2412 TAAAAA... 2434
 Db 2362 TTGCTACTTAAAAA... 2384

RESULT 15
 DB3075
 DEFINITION Lithospermum erythrorhizon mRNA for phenylalanine ammonia-lyase,
 complete cds. 2441 bp mRNA linear PLN 06-FEB-1999

ACCESSION D83075
 VERSION D83075.1 GI:2911121
 KEYWORDS phenylalanine ammonia-lyase,
 Lithospermum erythrorhizon dedifferentiated cells cell suspension
 cultures cell line:m18 cDNA to mRNA, clone:lib:expression library
 of Lithospermum cells cultured in pigment production medium
 clone:LEPAL-1.

ORGANISM Lithospermum erythrorhizon
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; euasterids I incertae sedis; Boraginaceae;
 Lithospermum.

REFERENCE 1 (sites)
 Yasaki, K., Kataoka, M., Honda, G., Severin, K. and Heide, L.
 cDNA cloning and gene expression of phenylalanine ammonia-lyase in
 Lithospermum erythrorhizon
 Biosci. Biotechnol. Biochem. 61 (12), 1995-2003 (1997)
 98101967

REFERENCE 2 (bases 1 to 2441)
 Yasaki, K.
 Direct Submission
 Submitted (16-Jan-1996) Kazufumi Yasaki, Graduate School of
 Agriculture, Kyoto University, Lab. Molecular & Cellular Biology,
 div. Applied Life Sciences, Kitashitakawa, Kyoto, Kyoto 606-01,
 Japan (E-mail:yasaki@kais.kyoto-u.ac.jp, Tel:075-753-6384,
 Fax:075-753-6398)

FEATURES
 source
 Location/Qualifiers
 1..2441
 /organism="Lithospermum erythrorhizon"
 /db_xref="taxon:34254"
 /clone="LEPAL-1"
 /cell_line="M18"
 /tissue_type="cell suspension cultures"
 /clone_lib="expression library of Lithospermum cells
 cultured in pigment production medium"
 /dev_stage="dedifferentiated cells"
 /note="Sieb. et Zucc."
 161..2293
 /codon_start=1
 /product="phenylalanine ammonia-lyase"
 /protein_id="BAA24928.1"
 /db_xref="GI:2911122"
 /translation="METIVENGCKTMEFCMKDPLNMEMASFSKMGSHLDEVKNVAE
 FRKPVVOLAGTLTIGVAAIAARDGVTELEAAEGYKASSDVMMSMNGTDSY
 GVTGFGATSHRTKOGALOKELIRFLNMGIPFNGTSTHPLHSATPAMVRINT
 LLOGSGIRRELEAITKELNNTITPCLPTGTTASGDLVPLSTAGLTGSPNKA
 VQPTCEKINAEARLAGISTGFELDPKRGALVNGTAGSMASWVLEANTLAVL
 SEVISAIFAEVWNGKPEPTDLHKLHHPGQIPAAIMHIIIDGSOYVAAOKLHEM
 DPLQPKODRYALRTSPQWIGPOLLEVIKSAIKMIEREINSVNPIIDVSRNALIGG
 NFOGTPICVANDNTRLIASIGKLLFAQSELDVYNNLPSNLGSRPSLDYEFK
 GAELIMASYCELOLAFNPVTHVQSAEONHODVNSGLISSKRTSEAVIILKMSF
 FLVLFQAVDRIHEENVRLAVKNTSDVAKRLLTTSVNELHSPSESDILLRVDRE
 EVFAYADPCITITTPIMOKURETLVGHALDNGNKKDVNTSIFPKIALIFERIKAIL
 PKVHNARASVENCIPATISNRIEGRSYPIKRVRELGTLLTGKRVSPGHEIDKY
 FTAKNEKLVDPILACTLAWNGAPLPIC"

polyA_signal
 BASE COUNT 663 a 526 c 567 g 685 t
 2293..2298

Query Match 50.3%, Score 1227.4, DB 8, length 2441;
 Best local Similarity 72.9%, Pred. No. 2,86-283;
 Matches 1595; Conservative 0; Mismatches 586; Indels 6; Gaps 1;

QY 135 AGCTAATGAGATGCTTAAATGATGTCATGAGATCATTAATGATGAGATGAG 194
 Db 180 ATGGAATGGAATAAACTATGAGATTTCATGAAAGATCATTAATGAGATGAGAT 239

QY 195 CGAGCGCTTGACCGGAGAGTCACTTGATGAGTGAAGAAATGCTTGGAGCTTGA 254
 Db 240 CTGAGTCAATGAAAGGAGACACTTGTATGAAATGAAATGATGAGATGAGTGA 299

QY 255 ACCCGTGTGAGAGCTTCGAGAGAGAGAGAGCTTACAGTTTCACATGGCGGATCG 314
 Db 300 AACCGGTGTGAGAGCTTCGAGAGAGAGAGAGCTTACAGTTTCACATGGCGGATCG 359

QY 315 CTGCTATGACAGTACACACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 374
 Db 360 CCGGTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 413

QY 375 AGCGAGTATGATTTGGCTTATGAGAGAGATGAATAAAGCACTGATGATGCTCA 434
 Db 414 AGCTAGTATGATTTGGCTTATGAGAGAGATGAATAAAGCACTGATGATGCTCA 473

QY 435 CCACCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
 Db 474 CCACCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 533

QY 495 AGCTATGATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 554
 Db 534 AACTATGATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 593

QY 555 TTGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614
 Db 594 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 653

QY 615 ACACCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
 Db 654 ATTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 713

QY 675 CCCCTGTTTACCTCCGCTGAGACATGACGCTCCGCTGATGATGATGATGATGAT 734
 Db 714 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773

QY 735 ACATGCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 794
 Db 774 ACATGCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 833

QY 795 TCCCTAATGCGGAG 854
 Db 834 ACATGCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 893

QY 855 ACACCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 914
 Db 894 ACATGCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 953

QY 915 TCGCTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974
 Db 954 TCGCTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1013

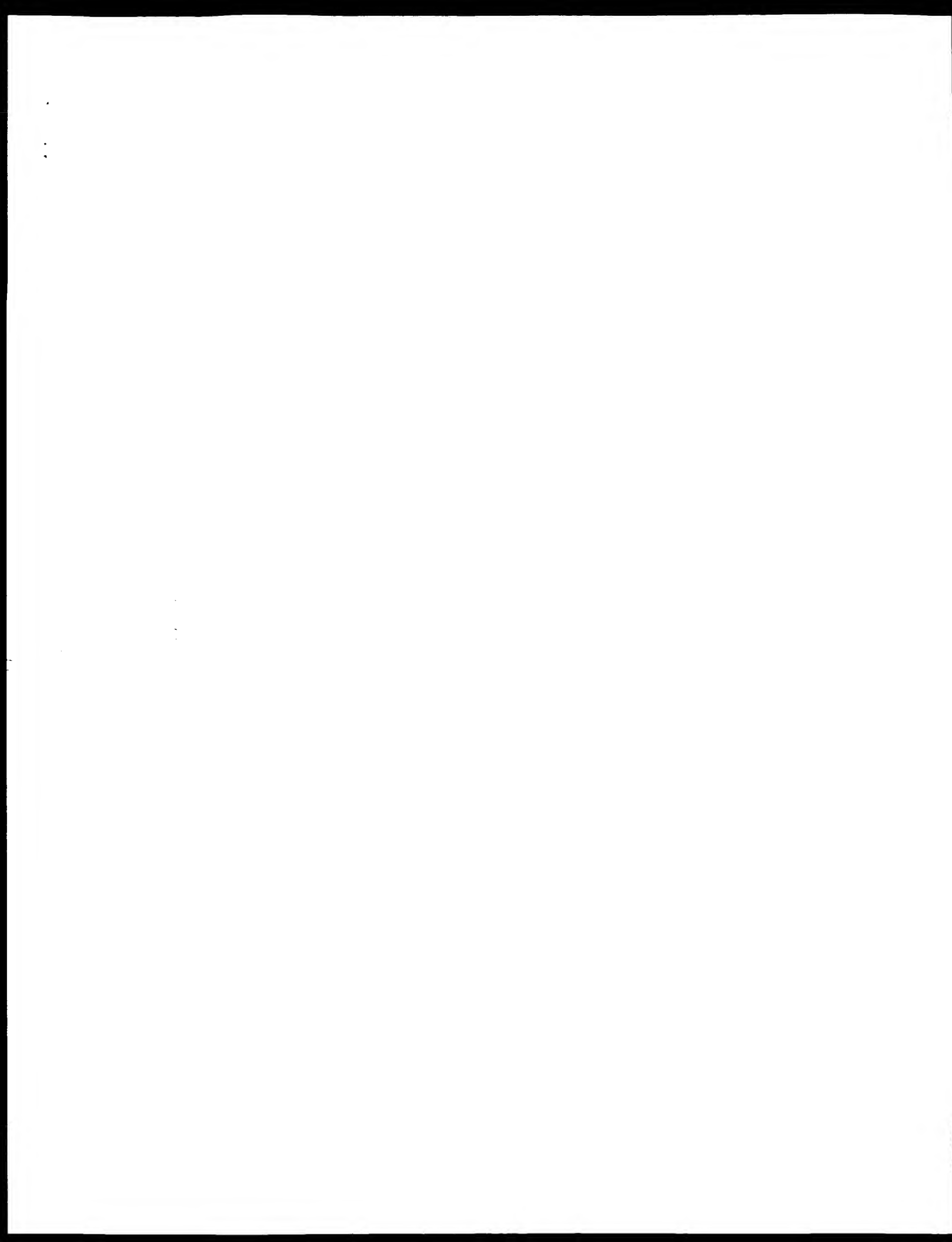
QY 975 CTGAGTATGAG 1034
 Db 1014 CTGAGTATGAG 1073

QY 1035 ACCCGTGTGAGAGCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1094
 Db 1074 ACCCGTGTGAGAGCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1133

QY 1095 TCAAGCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1154
 Db 1134 TCAAGCGGCTTGGCGGACACCTTCACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1193

QY 1155 ATGCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1214
 Db 1194 ATGCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253

QY 1215 CCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274



Genforce version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 28, 2003, 21:47:27 : Search time 3111 Seconds

(without alignments)
12712.754 Million cell updates/sec

Title: us-09-964-992a-3

Perfect score: 2442

Sequence: 1 gagcaatcagcaatcc aaaaaaaaaaaaaa 2442

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 800774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

EST:
1: em_estha:
2: em_esthm:
3: em_estin:
4: em_estmu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_estr:
9: qb_estl:
10: qb_est2:
11: qb_est3:
12: qb_est4:
13: qb_est5:
14: qb_est6:
15: em_estom:
16: em_estom:
17: qb_est:
18: em_estom:
19: em_estom:
20: em_estom:
21: em_estom:
22: em_estom:
23: em_estom:
24: em_estom:
25: em_estom:
26: em_estom:
27: em_estom:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	942.4	38.6	2598	11	AY103647 Zea mays
2	857.6	35.1	2299	11	AY106831 Zea mays
3	759	31.1	759	14	HO850732 Zea mays
4	758	31.0	1864	11	AY104679 Zea mays
5	715	29.3	741	14	BO941054 Q6F2IN06
6	708.2	29.0	770	14	FM009214 GCH9N08.7

7	694.4	28.4	710	14	BO869202
8	682.6	28.0	712	14	BO869215
9	661.4	27.1	679	14	BO998011
10	657.4	26.9	661	14	BO869819
11	656.8	26.9	679	14	BO007076
12	623	25.5	655	14	BO001313
13	564.2	23.1	674	14	BO865594
14	526.8	21.6	711	14	BO007176
15	523.8	21.4	696	14	BO024517
16	518.8	21.2	634	14	BO995720
17	518.8	21.2	802	14	BO645264
18	514.6	21.1	697	14	BO028965
19	513.6	21.0	707	14	BO028661
20	508.2	20.8	807	14	BO929414
21	507.2	20.8	707	14	BO028321
22	506.2	20.7	715	14	BO028594
23	500.6	20.5	692	14	BO029550
24	498.8	20.4	779	14	BO351058
25	496.8	20.3	713	14	BO027265
26	496.8	20.3	792	12	BO645267
27	494.6	20.3	799	12	BO887997
28	494	20.2	806	13	BM405502
29	491.4	20.1	687	14	BO028544
30	490.6	20.1	676	14	BO028141
31	490.4	20.1	508	14	BO987939
32	486.8	19.9	699	14	BO029105
33	476.8	19.5	819	12	BO003440
34	472.2	19.3	767	14	BO045673
35	471.6	19.3	479	14	BO995697
36	471.4	19.3	734	14	BO943329
37	471.2	19.3	624	10	AM218834
38	471	19.3	726	14	BO968511
39	469	19.2	631	14	BO003880
40	468.2	19.2	626	14	BO004978
41	467.4	19.1	613	14	BO995679
42	467.4	19.1	631	12	BO925579
43	465.6	19.1	794	12	BO647641
44	465.2	19.0	639	14	BO027438
45	464.6	19.0	744	12	BO597014

ALIGNMENTS

RESULT 1	AY103647	2598 bp	mRNA	linear	HTC 25 MAY 2002
LOCUS	AY103647				
DEFINITION	Zea mays PC0142084 mRNA sequence.				
ACCESSION	AY103647				
VERSION	AY103647.1	01-21-06-725			
KEYWORDS	HTC.				
SOURCE	Zea mays.				
ORGANISM	Zea mays.				
REFERENCE	1 (bases 1 to 2598)				
AUTHORS	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.				
TITLE	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 2598)				
AUTHORS	Cooper, R.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
FEATURES	Location/Qualifiers				
SOURCE	1..2598				
	/organism="Zea mays"				
	/db_xref="MaizeDB:638378"				
	/db_xref="taxon:4577"				
	/clone="PC0142084"				

```
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
```

assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overdo addressing of BACs in conjunction with the Maize Mapping Project."

BASE COUNT	490 a	907 c	795 g	406 t
ORIGIN				

Query Match	38.68;	Score 942.4;	DB 11;	Length 2598;
Best Local Similarity	66.68;	Pred. No. 9.2e-115;		
Matches 1383; Conservative	0;	Mismatches 686;	Indels 9;	Gaps 2;

QY	170	GATCCATTGAAAGTGGGAGTTGGACGGGAGGCGTTGACCGGAAAGTCACTTCATGAAGTC	223
Db	140	GACCCGTGTAACCTGGGGCGGGGGGGGAGACCTGGGAGAGGCTACCTGGACGAGGATG	199
QY	230	AGCAAGATGGTTGGCGAGTTCAGAAAGGCGGTGGTGAAGCTGGGAGGAGACGCTTACA	289
Db	200	AAGCGCATGGTGGCGGAGGGCGGGGAGCCCGTGGTCAATGATCGAGGGTTCACCCCTCCGC	259
QY	290	GTTTCTGAGTGGGGGAGTCCGAGCTGCTATGACAGTACACCGGTGAAGGTGGAGCTG	349
Db	260	GTCGACCAAGGTGGGCGGCTGCGCTCGCCACAAAGACGCTCGGGGTCGCGGTGACGCTC	319
QY	350	TCCGAAACCCCGAAGCGCTGGAGTTAAGGGCAGTAGTGAATTGGGTTATGAGACGATGAAT	409
Db	320	GACGAGGAGGCGCGCCCGCCGTCAGAGGCGAGGAGATGGATCCTGATGCTATCGC	379
QY	410	AAAGAACTGATATATTATTAATGTATACAGCGATTGGTGGTATGATATGAAATAT	469
Db	380	CACGGCGGCACATCTACGGCGTCACACCGCGCTTGGGGGCACTCCACCGCGCGAC	439
QY	470	AAGCAAGCGGATGTGTTTAAAGAAATGATCATTAATTTTAAAGCGCGAATATTGGG	529
Db	440	AAGGACGGGCGCGCGCTCGAGTGGAGTGTGCTAGGACATCTACAGCGCGAATCTTGGC	499
QY	530	AATGGAACGGAAACAAAGCCACACACTTTCACATTCAGCGCACAGAGCCCGCATGTATGTC	589
Db	500	ACCGGCGAGCA---CGGGTACAGCGCTGGCTGGAGAGTTATCTGGCGGCGATGCTGGTG	556
QY	590	AGAAATCAACCCCTCTCCAGGGTTACTCCGCGATCCGATGCGAATCTTGAAGACCATC	649
Db	557	CGCATCAACCCCTCTCCAGGGCTACTCCGGGATCCGCTTGAGATCTCTCAGGCGATC	616
QY	650	ACCAATGTTCTTAAATAAATATATATTTTAAATATATATATATATATATATATATAT	709
Db	617	ACGAAAGTGGCTCAACACCGGTTGACCGCCCTGCTTGGTATATATATATATATATATAT	676
QY	710	TCCGGTGACCTTGTCCCAT	769
Db	677	TCCGGTACCTTGTCCCGTCTCTCTATATATATATATATATATATATATATATATAT	736
QY	770	AAATGCTTGGCGCTAT	829
Db	737	CAGGCGCTCACCGTCAACGGGACGAAGTGAAGCGCCCGCCACCGCTTCAACATGCCCGC	796
QY	830	GTTGAAGTGGGTTCTTGATGATGACCGGAAAGAGGGCTAGACTTTTAAAGGACAC	889
Db	797	ATCAGAGGGCGGCTTCTTCAAGTCTCAACCGCCAAAGGATGGCTTGCATGTCAGGGCAG	856
QY	890	GCGCTGGGGTCCGGAGTGGCTTCATATGTTCTATATTTGATGCTAAATGTACTGCTTTTG	949
Db	857	TCCGTGGGGCTCCGCGCTCGCGGCCACCGGTATGATAGAGCGCCACGCTCTGACCTCTG	916
QY	950	TCCGAATGTTATTCGGGATCTCTCGTGAAGTATGACAGGAGACCGGATTTACCAT	1009
Db	917	TCCGAGTCTCTGTCGCGCT	976
QY	1010	CATTTGACACAAATTTGAAGTATATACCTGGTAAATTTGAAGTGGTGGATATATGAG	1059

[illegible]

University of California at Davis (UCD)
 Asmundson Hall, mcd, Davis, CA 95615, USA

Tel: 1 (530) 742-1742
 Fax: 1 (530) 752-9659

Email: akozik@ucdavis.edu [michelmore@ucdavis.edu]
 belongs to contig q9_CA.Contig196760, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: q9H9, row: N, column: 08.

FEATURES

Source

Location/Qualifiers

1..770

/organism="Lactuca sativa"

/cultivar="L. serriola"

/db_xref="taxon:4236"

/clone="OG9N08"

/clone_lib="OG_FPCBU lettuce serriola"

/lab_host="E.coli"

/note="Vector: pBRCDNA51AB; the library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformation made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdb.ucdavis.edu/>

BASE COUNT

188 a 211 c 200 g 171 t

ORIGIN

Query Match

Best Local Similarity 29.0%; Score 708.2; DB 14, Length 770,
 98 RW, Pred. No. 9, 9e 84,

Matches 724; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

390 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 449
 39 GGGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 98
 450 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 509
 99 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 158
 510 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 569
 159 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 218
 570 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 529
 219 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 278
 630 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 689
 279 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 338
 690 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 749
 339 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 398
 750 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 809
 399 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 458
 810 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 869
 459 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 518
 870 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 929
 519 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 578
 930 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 989
 579 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 638

990 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 1049

639 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 698

1050 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 1109

699 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 757

1110 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 1122

758 GCGTATGACGACATGAATGAAGCACTGATGATGCTGACGACGCTTGGCG 770

RESULT 7

H0869202

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

Location/Qualifiers

1..710

/organism="Lactuca sativa"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="OGD5H12"

/clone_lib="OG_AHCD1 lettuce salinas"

/lab_host="E.coli"

/note="Vector: pBRCDNA51AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformation made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdb.ucdavis.edu/>

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

TAG_LIB="OG_AHCD1 lettuce salinas"

```

|||||
DB 1 GTTGAAGTGGGTTCTTTCAGTACAGCCGAAAGAGAGGCTAGCACTTGTATTAAGCAGC 60
890 GCGGATGGGTTCCGATGGCTTCGATGGCTTCGATGGCTTCGATGGCTTCGATGGCTTCG
DB 61 GCGGATGGGTTCCGATGGCTTCGATGGCTTCGATGGCTTCGATGGCTTCGATGGCTTCG
950 TCGGAAGTGGTATCGGATGGCTTCGATGGCTTCGATGGCTTCGATGGCTTCGATGGCTTCG
DB 121 TCGGAAGTGGTATCGGATGGCTTCGATGGCTTCGATGGCTTCGATGGCTTCGATGGCTTCG
1010 CACTTGACACACAAATTTAAGCATTCCTGTTAAATCGAGGGGCGGCGATCATGAG 1069
DB 181 CACTTGACACACAAATTTAAGCATTCCTGTTAAATCGAGGGGCGGCGATCATGAG 240
1070 TATATTTGGACGGAACGATTCAGCTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
DB 241 TATATTTGGACGGAACGATTCAGCTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
99 TACAGAAACCTAAACAAAGATGGTATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
DB 301 TACAGAAACCTAAACAAAGATGGTATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
1190 TAAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
DB 361 TAAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
1250 GACAACCAATTCAGTACGCTTCGACGAAACAAAGCTTACAGCGGCTGATTCGACGAA 1309
DB 421 GACAACCAATTCAGTACGCTTCGACGAAACAAAGCTTACAGCGGCTGATTCGACGAA 480
99 ACCGCAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
DB 481 ACCGCAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
1370 ATGTTGCGTCAATTTCTGAGCTGGTAAAGATTTCTACAAACATGGATACCATGAA 1429
DB 541 ATGTTGCGTCAATTTCTGAGCTGGTAAAGATTTCTACAAACATGGATACCATGAA 600
99 CTCTCCGCGTGCAGTACCGTACGTTGACGAGGTTGCAAGGTTGAGAAATGCCGAG 1489
DB 601 CTCTCCGCGTGCAGTACCGTACGTTGACGAGGTTGCAAGGTTGAGAAATGCCGAG 660
99 GCTTCTACTGTTGAGCTTCATTTCTGCAATGATTCACGAAAC 1537
DB 661 GCTTCTACTGTTGAGCTTCATTTCTGCAATGATTCACGAAAC 707

```

```

RESULT 8
BO862125 712 bp mRNA linear EST 14-AUG-2002
DEFINITION OGC20H10.yg.ab1 OGC-ABCDI lettuce salinas Lactuca sativa cDNA clone
ACCESSION BO862125
VERSION BO862125.1 GI:22247670
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa.
REFERENCE 1 (bases 1 to 712)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellisson
P., Kolman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lal,Z.,
Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compenomics.ucdavis.edu/
unpublished (2002)
JOURNAL Contact: Alexander Kozik [R.W.Michelmore]
COMMENT Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)

```

```

Asmudson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@catv.ucdavis.edu [michelmore@vegmail.ucdavis.edu]
belongs to contig OGC-CA-Contig6760, see http://cgpdb.ucdavis.edu/
for details.
Plate: OGC20
FEATURES
source Location/Qualifiers
1..712
/organism="Lactuca sativa"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="OGC20H10"
/clone_11b="OG-ABCDI lettuce salinas"
/lb_host="E.coli"
/note="Vector: pBRCDMSf18: The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG-L18-OG-ABCDI lettuce salinas
TAG-TISSUE=chemical
TAG_SEQ=TGTACCCCGC"
BASE COUNT 189 a 173 c 174 g 176 t
ORIGIN
Query Match 28.08; Score 682.6; DB 14; Length 712.
Best Local Similarity 99.28; Pred. No. 2,3e-80;
Matches 707; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
830 GTTGAAGCTGGGTTCTTTCAGTACAGCCGAAACAAAGCGCTAGACATTTTACCGCACC 889
DB 1 GTTGAAGTGGGTTCTTTCAGTACAGCCGAAACAAAGCGCTAGACATTTTACCGCACC 60
890 GCGGTGGGTCGCGGAGTGGCTTCGATGTTCTATTTGATGCTAATGCTACTTGGCTTTG 949
DB 61 GCGGTGGGTCGCGGAGTGGCTTCGATGTTCTATTTGATGCTAATGCTACTTGGCTTTG 120
99 TGGGAATGATGCTTCGATGTTCTATTTGATGCTAATGCTACTTGGCTTTG 1009
DB 121 TCGGAAGTGGTATCGGCAATCTTCGAGGTATTCGAGGAGCGGAGGCTTTACCGAT 180
99 TCGGAAGTGGTATCGGCAATCTTCGAGGTATTCGAGGAGCGGAGGCTTTACCGAT 180
1010 CACTTGACACAAATTTGAAACATCACCGTGGTCAAAATGAGGCGGCGGATCATGAG 1069
DB 181 CACTTGACACAAATTTGAAACATCACCGTGGTCAAAATGAGGCGGCGGATCATGAG 240
99 TATATTTGGACGGAACGATTTACGTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
DB 241 TATATTTGGACGGAACGATTTACGTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
1130 TTAAGAAACCAAAACAAAGATGATGTTCTGCAATGATTCGAAATGCTTGAAGCT 1189
DB 301 TTAAGAAACCAAAACAAAGATGATGTTCTGCAATGATTCGAAATGCTTGAAGCT 360
99 TTAAGAAACCAAAACAAAGATGATGTTCTGCAATGATTCGAAATGCTTGAAGCT 360
1190 CAAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
DB 361 CAAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
1250 GACAACCAATTCAGTACGCTTCGACGAAACAAAGCTTACAGCGGCTGATTCGACGAA 1309
DB 421 GACAACCAATTCAGTACGCTTCGACGAAACAAAGCTTACAGCGGCTGATTCGACGAA 480
99 ACCGCAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
1310 ACCGCAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
DB 481 ACCGCAATGGAATGATATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATATCTT
1370 ATGTTGCGTCAATTTCTGAGCTGGTAAAGATTTCTACAAACATGGATACCATGAA 1429
DB 541 ATGTTGCGTCAATTTCTGAGCTGGTAAAGATTTCTACAAACATGGATACCATGAA 600

```


Email: akozik@atgc.org [michelmore@vegmil.ucdavis.edu]
 belongs to contig OG_CA_Contig6760, see http://cgpdb.ucdavis.edu/
 for details.

FEATURES
 source
 Location/Qualifiers

1..661
 /organism="Lactuca sativa"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="OG67601"
 /clone_lib="OG_ABCDI lettuce salinas"
 /lab_host="E.coli"
 /note="Vector: pBRCDNA5FIAB: The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG_LIB-OG_ABCDI lettuce salinas
 TAG_TISSUE="chemical induction
 TAG_SFO="TGAOCCGGG"

BASE COUNT 160 a 185 c 178 g 138 t

Query Match 26.9%; Score 657; DB 14; Length 661;
 Best Local Similarity 100.0%; Pred. No. 5,4e-77;

Matches 657; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OG 268 GCTCGGAGGAGAGACGCTTACAGTTCTCAGGTGGCGGGGATCCGACCTGCTAATGACAG 327
 |||||||
 Db 1 GCTCGGAGGAGAGAGCTTACAGTTCTCAGGTGGCGGGGATCCGACCTGCTAATGACAG 60
 |||||||
 OG 328 TGACACCGTGAAGTGGAGCTGTCGGAAGCCGCGGCTGAGCTGAAGCGGAGTAGTGA 387
 |||||||
 Db 61 TGACACCGTGAAGTGGAGCTGTCGGAAGCCGCGGCTGAGCTGAAGCGGAGTAGTGA 120
 |||||||
 OG 388 TTGGGTTATGAGAGCATGATTAAGAACTGATAGTATGATGATGATGATGATGATGATG 447
 |||||||
 Db 121 TTGGGTTATGAGAGCATGATTAAGAACTGATAGTATGATGATGATGATGATGATGATG 180
 |||||||
 OG 448 CCGGACCTTCACCGGAGACGTAACGAGCGGCTCTTACAGAGAGGAGCTCATTAAT 507
 |||||||
 Db 181 CCGGACCTTCACCGGAGACGTAACGAGCGGCTCTTACAGAGAGGAGCTCATTAAT 240
 |||||||
 OG 508 TTTCAGACCGCGGATATTCGCCAATGCAAGCAAGCAACACACATTCACATTCAGC 567
 |||||||
 Db 241 TTTCAGACCGCGGATATTCGCCAATGCAAGCAAGCAACACACATTCACATTCAGC 300
 |||||||
 OG 568 CATTCAGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 627
 |||||||
 Db 301 CACCAAGACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 |||||||
 OG 628 ATTTCAGATCTTGAAGGACGATCAGCAAGTTCCTTAACACACATTCACACCTTCAT 687
 |||||||
 Db 361 ATTTCAGATCTTGAAGGACGATCAGCAAGTTCCTTAACACACATTCACACCTTCAT 420
 |||||||
 OG 688 CCTCCCTGGAACCATCAGCCCTCCGCTGAGCTTGTCCCATTAATACATGCGCGCT 747
 |||||||
 Db 421 CCTCCCTGGAACCATCAGCCCTCCGCTGAGCTTGTCCCATTAATACATGCGCGCT 480
 |||||||
 OG 748 CTTAATCCGCGCGCTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 807
 |||||||
 Db 481 CTTAATCCGCGCGCTTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 540
 |||||||
 OG 808 AAAGGCTTCTGACCGCGAGTGAAGTGGGTTCTTCAGATTAACAGCGGAAGAAG 867
 |||||||
 Db 541 AAAGGCTTCTGACCGCGAGTGAAGTGGGTTCTTCAGATTAACAGCGGAAGAAG 600
 |||||||
 OG 868 GCTAGCATTTTAAAGGCGACCGCGTGGGCTCCGCGATGCTTCATGTTCTATT 924
 |||||||

Db 601 CTTAGACATCTTAAAGCGACCGCGGCGGCGGAGATGCTTCATGTTCTATT 657

RESULT 11
 LOCUS BU007619 679 bp mRNA linear EST 22-AUG-2002
 DEFINITION OG44h01.y9.ab1 OG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
 OG44h01.y9.ab1 OG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
 ACCESSION BU007619
 VERSION BU007619.1 GI:22442014
 KEYWORDS EST.
 ORGANISM Lactuca sativa.
 SOURCE Lactuca sativa.
 Lactuca sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
 Lactuca.

REFERENCE 1 (bases 1 to 679)

AUTHORS

Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Riesberg, L.,
 Liu, H., Van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison,
 P., Kolman, T., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,
 Church, S., Jackson, L. and Bradford, K.

Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compenomics.ucdavis.edu/
 Unpublished (2002)

JOURNAL

Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Ashmound Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmil.ucdavis.edu]
 belongs to contig OG_CA_Contig6760, see http://cgpdb.ucdavis.edu/
 for details.

Plate: OG44 row: h column: 01.

FEATURES

source

Location/Qualifiers

1..679
 /organism="Lactuca sativa"
 /cultivar="L.serriola"
 /db_xref="taxon:4236"
 /clone="OG44h01"
 /clone_lib="OG_EFGHJ lettuce serriola"
 /lab_host="E.coli"
 /note="Vector: pBRCDNA5FIAB: The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG_LIB-OG_EFGHJ lettuce serriola
 TAG_TISSUE="leaves dark grow
 TAG_SFO="GCTAGCGCGG"

BASE COUNT 190 a 145 c 186 g 158 t

Query Match 26.9%; Score 656.8; DB 14; Length 679;
 Best Local Similarity 98.8%; Pred. No. 5.6e-77;

Matches 672; Conservative 0; Mismatches 7; Indels 1; Caps 1;

OG 1 GAGCAATCTGATCATATACCATTCATGACGATCAAAAGATGATGATGATGATGATGATG 60
 |||||||
 Db 1 GAGCAATCTGATCATATACCATTCATGACGATCAAAAGATGATGATGATGATGATGATG 60
 |||||||
 OG 61 ACACATATGATGTTCTGTTCTTTCATGATATAGCTACAAATCTGTAATAAATAAT 120
 |||||||
 Db 61 ACACATATGATGTTCTGTTCTTTCATGATATAGCTACAAATCTGTAATAAATAAT 120
 |||||||
 OG 121 GGAGAACGGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 |||||||
 Db 121 GGAGAACGGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 |||||||

QY 181 CTGGGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 240
 |||||||
 Db 181 CTGGGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 240
 QY 241 TGGCGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 300
 |||||||
 Db 241 TGGCGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 300
 QY 301 GATGGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 360
 |||||||
 Db 301 GATGGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 360
 QY 361 GAGGCGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 420
 |||||||
 Db 361 GAGGCGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 420
 QY 421 TACGTTATGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 480
 |||||||
 Db 421 TACGTTATGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 480
 QY 481 TGCTTTATGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 540
 |||||||
 Db 481 TGCTTTATGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 540
 QY 541 AACAGCGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 600
 |||||||
 Db 541 AACAGCGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 600
 QY 601 CCGCTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 660
 |||||||
 Db 601 CCGCTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 660
 QY 661 TACGACGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 720
 |||||||
 Db 661 TACGACGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 720

RESULT 12
 B0001313 655 bp mRNA linear EST 23-Aug-2002
 LOCUS OGG27H21.yr.abi OG_EFGHU lettuce serriola lactuca sativa cDNA clone
 DEFINITION OGG27H21, mRNA sequence.

ACCESSION B0001313
 VERSION B0001313.1 GI:22435708
 KEYWORDS EST.
 SOURCE Lactuca sativa.
 ORGANISM Lactuca sativa.
 Eukaryota; Viridiplantae; Stratiolypeta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
 Lactuca.

REFERENCE 1 (bases 1 to 655)
 Kozik, A., Michelmore, R. W., Knapp, S., Matvienko, M., Riesberg, L.,
 Liu, H., Van Damme, M., Javelle, D., Chevalier, P., Ziegler, J., Ellison,
 P., Kolman, J., Staubach, M. S., Livingston, K., Zhou, Y., Lai, Z.,
 Church, S., Jackson, L. and Bradford, K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compenomics.ucdavis.edu/
 Unpublished (2002)

TITLE JOURNAL
 COMMENT Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Assamundson Hall, UCD, Davis, CA 95615, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
 belongs to contig OG_CA_Config6760, see http://cgdb.ucdavis.edu/
 for details.

FEATURES
 source
 Plate: OGG27 row. H column: 21.
 1. 655
 Location/Qualifiers
 /organism="Lactuca sativa"

/cultivar="L.serriola"
 /db_xref="taxon:4236"
 /clone="OGG27H21"
 /clone_lib="UG_EFGHU lettuce serriola"
 /lab_host="E.coli"
 /note="Vector: pBR322/SHV; The library was constructed
 from 10 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at http://cgdb.ucdavis.edu/
 TAG-SEQ-Not found"

BASE COUNT 175 a 159 c 156 t
 ORIGIN

Query Match 25.5%, Score 623; DH 14; Length 655;
 Best Local Similarity 96.9%; Pred. No. 1,50-72;
 Matches 635; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 820 TGAGGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 879
 |||||||
 Db 1 TGGAGGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 879
 QY 880 TAAAGGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 939
 |||||||
 Db 61 TAAAGGAGTTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 939
 QY 940 TCGCTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1000
 |||||||
 Db 121 TCGCTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1000
 QY 1000 GTTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1059
 |||||||
 Db 181 GTTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1059
 QY 1060 GATTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1119
 |||||||
 Db 241 GATTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1119
 QY 1120 AATGAGCGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1179
 |||||||
 Db 361 AATGAGCGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1179
 QY 1180 GATTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1239
 |||||||
 Db 361 GATTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1239
 QY 1240 TCGCTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1299
 |||||||
 Db 421 TCGCTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1299
 QY 1300 GTTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1359
 |||||||
 Db 481 GTTTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1359
 QY 1360 GATTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1419
 |||||||
 Db 541 GATTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1419
 QY 1420 AATTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1479
 |||||||
 Db 601 AATTCAGCTTCAGACGCGGCTTGACCGGAGCTTCAGCTTGATGAGCTGACAGATGCT 1479

RESULT 13
 B0865594 674 bp mRNA linear EST 14-Mar-2002
 LOCUS OGG5F22.yr.abi OG_EFGHU lettuce serriola lactuca sativa cDNA clone
 DEFINITION OGG5F22, mRNA sequence.
 ACCESSION B0865594

VERSION	KEYWORDS	SOURCE	ORGANISM
H086594.1	GI:2251059	Lactuca salvia.	Lactuca salvia.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots; Asteridae; euasterids II; Asterales; Astuceaceae; Lactuceae; Lactuca.			
1 (bases 1 to 674)			
Kozik,A., Michelmore,R.W., Knapp,S., Melvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Jai,T., Church,S., Jackson,U., and Bradford,K.			
Lettuce and Sunflower ESTs from the Compositae Genome Project			
http://compgeonomics.ucdavis.edu/unpublished (2002)			
Contact Alexander Kozik [R W Michelmore]			
Department of Vegetable Crops, R.W.Michelmore lab			
University of California at Davis (UCD)			
Asmundson Hall, UCD, Davis, CA 95616, USA			
Tel.: 1-(530)-742-1742			
Fax: 1-(530)-752-9659			
Email: akozik@ucdavis.org [michelmore@veemail.ucdavis.edu]			
belongs to contig QC_CA_Contig517, see http://cypdb.ucdavis.edu/details.			
Plate: QCC5 row: F column: 22.			
Location/Qualifiers			
1..674			
/organism="Lactuca salvia"			
/cultivar="Salinas"			
/db_xref="taxon:4236"			
/clone="QCC5P2"			
/clone.lib="UG ABCDI Lettuce salinas"			
/lab.host="E.coli"			
/note="vector: pRCMDAS1AB: The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cypdb.ucdavis.edu/TAGLib-QC_ABDI lettuce salinas			
TAG_Lib-QC_ABDI lettuce salinas			
TAG_TISSUE=chemical induction			
TAG_SEQ=TGACGCCGG*			
BASE COUNT	167 a	179 c	156 t
ORIGIN			
Query Match	23.1%	Score 564.2,	DB 14, Length 674;
Best Local Similarity	89.9%	Pred. No 6.9e-65;	
Matches 603; Conservative	0;	Mismatches 68; Indels	0; Gaps 0
Y	371 GTTAAGCGAGTATGATTGGCTTAATGCAGACATGAATAAGCAACTGATTATGT	430	
D			
D	2 GTCAAGCGAGTAGTATGGCTTCATGAGACTATGAAATAAGCAAGCATGCTATGT	61	
Y	431 GTCAACACCGCGTTGGCGGCGCACCTGTCTACATGGAATAAAGCAAGCGTGCTTACAG	490	
D			
D	62 GTCACTACCGCGTTGGCGACTACCCCTACCCGACAAGAAAGAAGTGTCCTCTTCC	121	
Y	491 AAGGAGCTATTCATTTTGAACCGCGGAATATTCGCGCAAATGCAGCAACAAACAGCAG	550	
D			
D	122 AAGGAGCTATTCATTTTGAACCGCGGAATATTCGCGCAAATGCAGCAACAAACAGCAG	181	
Y	551 ACACATTCACATTCACGCCACAGCGCGCATGATGTCAGAAATCAACACCCTCTCCAG	610	
D			
D	182 ACACTTCACATTCACGCCACAGCGCGCATGTCAGAAATCAACACCCTCTTCA	241	
Y	611 GGTTACTCGCGCATCGATTGGAATCTTGGAAAGCATCAATCAAGTTCCTTAACATAC	670	
D			
D	242 GGTACTCGCGCATCGATTGGAATCTTGGAAAGCATCAATCAAGTTCCTTAACATAC	301	

[illegible]

HASH: COUNT	ORIGIN
169 a	192 c
194 g	152 t
4 others	

21.6%. Score 526.8; DR 14; Length 711;
ity 84.2%; Pred. No. 5e-60;
servative 0, Mismatches 111, Indels 0;

[illegible]

636 bp mRNA 112941 EST 23 AUG 2004
 CH-EPGCH sunflower RHA280 Helianthus annuus cDNA
 FF4607, mRNA sequence.
 .1 GI:22460937
 sunflower.
 us annuus
 Vitidiplantae, Streptophyta, Embryophyta; Tracheophyta;

Species: Magnoliophyta, eudicotyledons, core eudicots:
Asterales; eumastids II; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
I (bases 1 to 696)
Kozlik, A., Michelmore, R. W., Knapp, S., Malvienko, M., Kieseberg, T.,
Liu, H., Van Damme, M., Lavarelle, D., Chevalier, P., Ziegler, J., Ellison,
P., Kolman, J.J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lal, Z.,
Church, S., Jackson, L., and Bradford, K.
Letture and Sunflower ESTs from the Compositae genome project
<http://compositomus.ucdavis.edu/>
Unpublished (2002)
Contact: Alexander Kozik (R.W.Michelmore)
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundo Hall, UCD, Davis, CA 95616, USA
Tel.: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [mailto:moreveg@mail.ucdavis.edu]
belongs to contig QH_CA_Contig3142, see <http://cgpdb.ucdavis.edu/>
for details.
Plate: qHP4 row: e column: 07.
Location/Qualifiers
1..696
/organism="Helianthus annuus"
/cultivar="RNA280"
/db_xref="taxon:4232"
/clone="QH4e07"
/clone_lib="OH-EFGH sunflower RNA280"
/lab_host="E.coli"
/note="Vector: pBheNDASHAB. The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at <http://cgpdb.ucdavis.edu/>
TAG_LIB=QH-EFGH sunflower RNA280
TAG_TISSUE=chemical induction
TAG_SEQ=TGTAAGCGGG*

175 a 206 c 171 g 144 t

	21.4%	Score 523.8;	DB 14;	Length 696;
ch	1 Similarity	84.6%;	Prot No 1 3e-59;	
5588;	Conservative	0;	Mismatch 107;	Indels 0; gaps 0;

ATTTTCGGCAATGGAACGAAACAAGGCACACTTCCATTCCATCAGCACACAGCCGCC 580
ATTGTGGGTAAAGGGTAATATTTAGTATCATTAATAATTAATTAATTAATTAATTAATTA 60
TGATGTAT 640
TTCTTGGTCAGATCAACACACCCTCTCCAGAGAAATTCGGCAATATTCGAATCTTG 120
AAGGATATATTAAT 700
TACAGGATATATTAAT 180
TACAGGATATATTAAT 760
TACAGGATATATTAAT 240
TACAGGATATATTAAT 820
TACAGGATATATTAAT 880
TACAGGATATATTAAT 940
TACAGGATATATTAAT

```
Db 361 AATGATACCGGCGTGGGCTGCGGATGAGCTGATGTTGTTTGAAGCCAAATGTTCTT 420
QY 941 GCGTTCCTTCGCGAAGTGTATCGGCCATCTTCGCTGAGGTTATGCAAGGAAACCCGAG 1000
Db 421 GCATTGTTGTGGAGTGTGTTGCTGATTTTCAAGAAATATATGCAAGGAAAGCGGAG 480
QY 1001 TTACCGATCACTTGACACACAAATTAAGCATCACTGCTGATTAATTAAGGAGTGG 1060
Db 481 TTACCGAACATTTGACCCACAACTTAAGCACCACCCGCGCAATATAGAGCCGCGAGCC 540
QY 1061 ATCATGAGATATATTTTGAAGGAAAGGATTACGTCAAAGGCGCGCAAAAGTCCAGAA 1120
Db 541 ATCATGAGATATATTTTGAAGGAAAGTACTATGTAAAGCCGCGCAAAAGTCCATGAG 600
QY 1121 ATGAGCCGTTACAGAAACCAAAACAAATGCTATGCTGCTGCTGCTGCTGCTGCTG 1180
Db 601 ATGAGCCGTTACAGAAACCAAAACAAATGCTATGCTGCTGCTGCTGCTGCTGCTG 660
QY 1181 CTGGACCTCAATATGAAATATCGATCATCAAC 1215
Db 661 CTGGACCAAAATGAAATGATCATTCGTCGCCAC 695
```

Search completed: March 29, 2003, 00:34:00
Job time : 3134 secs

61 ACACGATTA;ATTGTTCT
27 |||||

Db 61 ACACAAATAGATGTCTCTTCTTCTTATGATCTAATGCTACAACTGTAATATATAT 120
QY 121 GGAGAACGGTAATCAGCTTAATGAGCTGCTTAATGAGTTGTCATCAAGATCCATTGAA 180
Db 121 GGAGAACGGTAATCAGCTTAATGAGCTGCTTAATGAGTTGTCATCAAGATCCATTGAA 180
QY 181 CTGGGGAGTTGAGGAGAGGCTTTGAGGAGAGTACCTTATGAGTGAAGAGATGCT 240
Db 181 CTGGGGAGTTGAGGAGAGGCTTTGAGGAGAGTACCTTATGAGTGAAGAGATGCT 240
QY 241 TCGGAGATTGAGAAACCGGCTGTGAGAGTGGAGAGACCGCTTACAGTTTTCAGCT 300
Db 241 TCGGAGATTGAGAAACCGGCTGTGAGAGTGGAGAGACCGCTTACAGTTTTCAGCT 300
QY 301 GCGGCGGATGCGAGCTGCTTAATGAGAGTACCGCTGAGAGTGGAGCTGTGAGAGCGG 360
Db 301 GCGGCGGATGCGAGCTGCTTAATGAGAGTACCGCTGAGAGTGGAGCTGTGAGAGCGG 360
QY 361 GAAGGCTGAGATTAAAGGAGTAGTATGAGTTATGAGAGCATGAAATAAAGAACTGA 420
Db 361 GAGGCGCTGAGATTAAAGGAGTAGTATGAGTTATGAGAGCATGAAATAAAGAACTGA 420
QY 421 TAGTATAGTGTGACAGCGGCTGTGAGGCGGACGCTGTGAGAGAGTAAATACCAAGCGG 480
Db 421 TAGTATAGTGTGACAGCGGCTGTGAGGCGGACGCTGTGAGAGAGTAAATACCAAGCGG 480
QY 481 TCGTTTACAGAGAGCTGATTAATGATTTTGAACCGCGGAAATATTTGGCAATCGAAGCA 540
Db 481 TCGTTTACAGAGAGCTGATTAATGATTTTGAACCGCGGAAATATTTGGCAATCGAAGCA 540
QY 541 AACAGAGCATATCTTCTTATTTAGCTTCTGAGAGCGGCGATGATGTCAGATTAACAC 600
Db 541 AACAGAGCATATCTTCTTATTTAGCTTCTGAGAGCGGCGATGATGTCAGATTAACAC 600
QY 601 CCTGCTGAGAGTGTACCTGCGGACCTGCGATTCGAGATCTTGAAGCGATCACCAATGCT 660
Db 601 CCTGCTGAGAGTGTACCTGCGGACCTGCGATTCGAGATCTTGAAGCGATCACCAATGCT 660
QY 661 TAACAAACATCACCCTTTTTCACCTCCCTGCGGAAACATCACCGCTTCCGCTGACT 720
Db 661 TAACAAACATCACCCTTTTTCACCTCCCTGCGGAAACATCACCGCTTCCGCTGACT 720
QY 721 TGTGCTATTTATCATCATGAGAGGCTTTTAACTGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 TGTGCTATTTATCATCATGAGAGGCTTTTAACTGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CCCCACCGGAGAGTCTCAATCCGAAAAAGGCTTCCGTCAGCGGAGTTGAGAGTGC 840
Db 781 CCCCACCGGAGAGTCTCAATCCGAAAAAGGCTTCCGTCAGCGGAGTTGAGAGTGC 840
QY 841 GTTCTTGAGATTACAGCCCAAGAGAGGCTAGACACTTGTAAAGCGGACCGCTGAGGCT 900
Db 841 GTTCTTGAGATTACAGCCCAAGAGAGGCTAGACACTTGTAAAGCGGACCGCTGAGGCT 900
QY 901 CCGGAGATGCTTCAATGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 960
Db 901 CCGGAGATGCTTCAATGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 960
QY 961 ATCGGCGATCTTCCGAGAGTATGCAAGGAAAGGAGGCTTACCGATCACCTTACACA 1020
Db 961 ATCGGCGATCTTCCGAGAGTATGCAAGGAAAGGAGGCTTACCGATCACCTTACACA 1020
QY 1021 CAATTTGAACATCACCCCTGCTCAAAATCGAGCGCGCGGATCATGAGATATTTTGA 1080
Db 1021 CAATTTGAACATCACCCCTGCTCAAAATCGAGCGCGCGGATCATGAGATATTTTGA 1080
QY 1081 CAGAGAGATTAGTAAAG 1140
Db 1081 CAGAGAGATTAGTAAAG 1140
QY 1141 AAAAGAGAGATGTTATGCTTCCCTTACATGCTCCCAATGCTGAGAGCTCAAAATGAG 1200
Db 1141 AAAAGAGAGATGTTATGCTTCCCTTACATGCTCCCAATGCTGAGAGCTCAAAATGAG 1200

QY 1201 AATCGATCATCAACCAATGATGAGAGGAAATCAATTCGGTCAATCAATCCATT 1260
Db 1201 AATCGATCATCAACCAATGATGAGAGGAAATCAATTCGGTCAATCAATCCATT 1260
QY 1261 GATGAGAGTTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 GATGAGAGTTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 AGTTTCATGAGACACACCCCTGCTGCAATTCGCAATTCGCAATTCGCAATTCGCA 1380
Db 1321 AGTTTCATGAGACACACCCCTGCTGCAATTCGCAATTCGCAATTCGCAATTCGCA 1380
QY 1381 ATTTCTGAGCTGAGTTAAAGATTTTCAACACATGATTTACATGAAATCTCCGCTG 1440
Db 1381 ATTTCTGAGCTGAGTTAAAGATTTTCAACACATGATTTACATGAAATCTCCGCTG 1440
QY 1441 AGCTAACCTGATTTGAGTACGAGGTTCAAGAGTGAAGAAATCGCATGAGCTTACTG 1500
Db 1441 AGCTAACCTGATTTGAGTACGAGGTTCAAGAGTGAAGAAATCGCATGAGCTTACTG 1500
QY 1501 TTTGAGCTTCAAGTTTCTGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCA 1560
Db 1501 TTTGAGCTTCAAGTTTCTGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCA 1560
QY 1561 CAATCAAGAGCTTAATTTCTGAGATTAAATTTCAAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 CAATCAAGAGCTTAATTTCTGAGATTAAATTTCAAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 CATCTTAAACATCATGCTGCTGACATACTTACTGCTCTATGCAATTCATCATGATT 1680
Db 1621 CATCTTAAACATCATGCTGCTGACATACTTACTGCTCTATGCAATTCATCATGATT 1680
QY 1681 CCATTTGAGAGAGAGATTAATTCAGACAGTGAAGAACCCGTAAGTCCGCAAAA 1740
Db 1681 CCATTTGAGAGAGAGATTAATTCAGACAGTGAAGAACCCGTAAGTCCGCAAAA 1740
QY 1741 GGTCTGACCATGAGGCTGACAGGAGCTCCAGCGTGCAGATTTCTGAGAAAGATCT 1800
Db 1741 GGTCTGACCATGAGGCTGACAGGAGCTCCAGCGTGCAGATTTCTGAGAAAGATCT 1800
QY 1801 CCTCGGTTGTTGATGCTGATGATGCTTCTGCTTCAATTCAGACAGGTTTCACTGAC 1860
Db 1801 CCTCGGTTGTTGATGCTGATGATGCTTCTGCTTCAATTCAGACAGGTTTCACTGAC 1860
QY 1861 ATACCATTAATGAGAGAGCTCGAGAGGTTCTGCTGAGCAGCAGCTTAACCAAGGCG 1920
Db 1861 ATACCATTAATGAGAGAGCTCGAGAGGTTCTGCTGAGCAGCAGCTTAACCAAGGCG 1920
QY 1921 AAGGAG 1980
Db 1921 AAGGAG 1980
QY 1981 GAAG 2040
Db 1981 GAAG 2040
QY 2041 GTGATTTCAACAGAGATTAAAGCTTGCAGATGTAACCGCTTGTATAGCTTTGAAAGGA 2100
Db 2041 GTGATTTCAACAGAGATTAAAGCTTGCAGATGTAACCGCTTGTATAGCTTTGAAAGGA 2100
QY 2101 GAG 2160
Db 2101 GAG 2160
QY 2161 CAGGAG 2220
Db 2161 CAGGAG 2220
QY 2221 AGGCTGAG 2280
Db 2221 AGGCTGAG 2280

[illegible]

```

1 FILE REFERENCE: 29-4 (PABA-0130)PV
2 CURRENT APPLICATION NUMBER: US/09/770,149
3 CURRENT FILING DATE: 2001-01-26
4 PRIOR APPLICATION NUMBER: 60/178,506
5 PRIOR FILING DATE: 2000-01-27
6 NUMBER OF SEQ ID NOS: 999
7 SOFTWARE: FASTSEQ for Windows Version 4.0
8 SEQ ID NO 476
9 LENGTH: 660
10 TYPE: DNA
11 ORGANISM: Arabidopsis thaliana
12 US-09-770-149-476
13
14 Query Match 11.1% Score 270.6; Dh 10; Length 660;
15 Best Local Similarity 72.4%; Freq. No. 2.9e-56;
16 Matches 351; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
17
18 Oy 1772 CACCGTCGAGATCTTCGCGAGCAAGAACATCCGCGCGCTTGTTGATGCTGAAATACGCTTC 1831
19 Db 659 CACGCTTCGCTGCTTCGCGCAAAAAGCATTTACTGCAAAAGTTTAGACCGCTGCAACAGCTCAAC 660
20
21 Oy 1842 GTTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1891
22 Db 599 ACATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 540
23
24 Oy 1892 CTCGTCGACGACCGCTCTAAACCAACGCGCAACGAGAGACACATACACCGCTGCATCTTC 1951
25 Db 539 ATTGTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 480
26
27 Oy 1952 CAAAGATGCTTACGCTTCGACAGACAGATTTGAAAGTCTGTTACCGCAAGAGATTCAGAGCT 2011
28 Db 479 CATAAGATATGACCTTTTCGAGAGAGAGGCTTAAAGGACAGTGCATCCGCAAGAGAGAGAGAG 420
29
30 Oy 2012 GTTGAATATGCTTATATATATATATATATATATATATATATATATATATATATATATATATAT 2071
31 Db 419 GCAAGACAGAGCTTACGATTAACGAAATATGCGTAATGCGTAATGCGTAATGCGTAATGCGTAAT 360
32
33 Oy 2072 TCGTACCGCTTGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2131
34 Db 359 TCGTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 300
35
36 Oy 2132 AAGTCGACGCTGCGCGGACAGAGAGCTTCGACAGAGCTTCGACAGAGCTTCGACAGAGCTTCG 2191
37 Db 289 AAGTCGACGCTGCGCGGACAGAGAGCTTCGACAGAGCTTCGACAGAGCTTCGACAGAGCTTCG 240
38
39 Oy 2192 ATTATTCATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2251
40 Db 239 ATCATTCATCCGATGATATATATATATATATATATATATATATATATATATATATATATAT 180
41
42 Oy 2252 TAGCA 2256
43 Db 179 TAGCA 175
44
45 RESULT 7
46 US-09-939-408A-29
47 Sequence 29, Application US/09/939,408A
48 Patent No. US20020102712A1
49 GENERAL INFORMATION:
50 APPLICANT: Yoshida, Roberta
51 TITLE OF INVENTION: Phenylation Ammonia Lyase Polypeptide and
52 TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
53 TITLE OF INVENTION: Using Same
54 FILE REFERENCE: 29479/COUSCA
55 CURRENT APPLICATION NUMBER: US/09/939,408A
56 CURRENT FILING DATE: 2001-08-24
57 PRIOR APPLICATION NUMBER: US 09/624,693
58 PRIOR FILING DATE: 2000-07-24
59 PRIOR APPLICATION NUMBER: PCT/US01/23270
60 PRIOR FILING DATE: 2001-07-24
61 NUMBER OF SEQ ID NOS: 30
62 SOFTWARE: PatentIn Ver. 2.0

```



```
Db 1253 TCGCCAAACACATGAGAGAGACAGCGCTGCGCCGACAGATCCGCAACGCTCAACTTCA 1312
QY 1377 CTCAATTTTGTGAGTGTGTTAGATTTTACAAACATGGATTTACCATGAACTCTCCG 1436
Db 1313 CGAGGCTCAGCGGATGCTCAACCGCCGATGACCCGCGCTCCGCTCCGCGCG 1372
QY 1437 GTGAGGTAACCCCTACTTTGAGCTACGGGTGAAAGTGGAGAAATCCGCAATGCTTCT 1496
Db 1373 CGGAA--GACCCCTGCTCTCTTACCATTTGAAGGCGCTGACATGCTGCGTGGGCGT 1429
QY 1497 ACTGTGCTGAGCTTCACTTTCGCAAAATCCAGTCAACCACTCTTCAAAACCGCGAC 1556
Db 1430 ACACGCTGCGAGTGGACACCTCCGCAACCCGTGACGACGATCTCCAGCGCGCTGAGA 1489
QY 1557 AACCAATCAAGAGTTAATCTCTGGGATTAATTTCAAGAGAAACCGAGAGGAG 1616
Db 1490 TGGCGAAACGAGCGGCTCAACCTCGCTTGGCTCATCTCGCTGCTCCACGACCGATCCA 1549
QY 1617 TCGACATCTTAAACTCATGCTGCTGACATTAATAGTCGCTCTATGCGCAATTCAT 1676
Db 1550 ACAGGCTCTTCTCTCTCTGCTGACACGACCTTATTTGCTTCTTCAAGTCATGACT 1609
QY 1677 TACGC 1681
Db 1610 TCGCG 1614
```

```
RESULT 15
US-09-878-574-3252
; Sequence 3252; Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3252
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: L1B3028-013-Q1-H1-C3
US-09-878-574-3252
```

```
Query Match 8.1%; Score 198.8; DB 10; Length 363;
Best Local Similarity 71.8%; Pred. No. 7.3e-39;
Matches 260; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
```

```
QY 1418 TTACCAATCAATCTCCGCTGACGTAAACCTAGTTGAGTACCGGTTCAAGGTGCA 1477
Db 2 TTGCTTTCAATCTCACTGCTCCAGCAAAACCCAGCTTGATTAAGATTCAGGAGCT 61
QY 1478 GAAATGCCCATGCTTCTACTCTCTGAGCTTCACTTTCGCAAAATCCAGTCAAC 1537
Db 62 GAAATTTGCGCATTTGATTTCTGTAATTTGCAATTTGGCGAATCCGCTGCAAGC 121
QY 1538 CATGTTCAAGCGCGGACACACATCAAGAGCTTAATTTCTCTGGATTAAATTCAGCG 1597
Db 122 CAGCTCAAAAGCGCGGACACACATCAAGATGTAACCTCTGCGGCTGATTTCAATCA 181
QY 1598 AGCAAAACCGGACGATCGACATCTTAAATCAATGATGCTGACATTAATTTAGTGGCT 1657
Db 182 AGCAAAATCTATGAGATTTATTTGAAATCTTAAATCTATGATTTCTGATGCTGCT 241
QY 1658 CTATGCCAATTCATGATTTACCGCATTTGAGAGAGAACATGAATTCAGAGTGAAGAC 1717
```

```
Db 242 CTTTGGCAAGCCATTGACTTTGAGGCAATTTGAGAGAGAAATTTGAAGAACACGCTCAAGAC 301
QY 1718 ACCGTAAATTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1777
Db 302 GTTGTAGTCAAGTTGCTTAAGAGACTCTTCAACCAAGCTGTAATGAGAGACTTCAACCT 361
QY 1778 TC 1779
Db 362 TC 363
```

Search completed: March 29, 2003, 02:08:56
Job time: 213 secs

1,2 485 AACGACCGAATAGTAAAGGAGTAAAGCGGCTTCGGCGGCACTTCTCAGCGGAGGAGCG 544


```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,000
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000,1003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
SEQUENCE CHARACTERISTICS:
LENGTH: 624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

Query Match	9.68	Score 235.2	DB 2	Length 624
Best Local Similarity	69.78	Pred. No. 3.7e-50		
Matches 318; Conservative	0	Mismatches 138	Indels	0
			Gaps	0

QY	845	TTCCAGTTACAGCCCAAGAAAGAAAGAGTATGACATTTGTTAAACGAGCCGCTGAGGCGG	904
Db	169	TTTAAATGCAAGCCCAAGGAAGACATGGCTCTCTGTCAACGGCACAACGGTGGGATCCGCC	228
QY	905	ATGGCTTCATGGTCTCAATTATGATGATGATCTTGCATGTGTGCGAAGCTGTATCG	964
Db	229	GTGGCCCGCTCCGCTCTGTGTATGACCAACATGTCTGGGCTGTGGCTGATGTTCTGCT	288
QY	965	GGCATCTTCGCTGAGGTTATGCAAGGGAAGCCGGAGTTATCCGATCACTGTGACACAAA	1024
Db	289	GGGCTCTTTCGCAAGTGTATGCAGAGGGAACCGGAGTTGTGATTCGTTAACCAACAG	348
QY	1025	TTGAAAGCATCACCTTGGTCAATCGAAGCGGCGGCGATCATGAGTATTTTGGACGA	1084
Db	349	TTCAAGCAACCAACCAAGCCGATCGAAGCCGCCGCCGCTCATGAGTTCTCTCTGACGCT	408
QY	1085	AGGCAATTACTCAAGGAGGAGGCAAAAAGGTCACAGAAATGACCCGTTACGAAACCAAA	1144
Db	409	AGGCACCTACGTGAAAGAGAGAGAGTGGCTTTATATGAAAGATATGTTATGATGAA	468
QY	1145	CAAGATCGTATCTCTCCGTACATCTCCCAATGGCTCGCACTCAAAATCGAATATTC	1204
Db	469	CAAGATCTGATATCTCTTCTGTAATATATGATATATGATGTTGGAGCTTCGATCTAAATCATC	528
QY	1205	CGATCATCAACCAAAATGATCGAGAGGGAATCATTCGTCACAGCAACCCATTGTC	1264
Db	529	CGCGCTCTCACTCACTCAACGACGGGAGATCAAAATTCGTCACAGCAATCCGTTATTC	588
QY	1265	GAGCTTCGGAAGCAAAAGTTTACCGGTGGTATC	1300
Db	589	GATCTCTCAGGAGCATGGCTGTCCAGCGGCGCAAC	624

RESULT 12
 US-08-975-316-9
 : Sequence 9, Application US/08975316
 : Patent No. 5952486
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: BLOKSBERG, Leonard N., HAYUKKALA, Ilkka
 : APPLICANT: and GRIERSON, Alastair W.
 :
 : TITLE OF INVENTION: MATERIALS AND METHODS FOR
 :
 : TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
 :
 : NUMBER OF SEQUENCES: 88
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Law Offices of Ann W. Speckman
 : STREET: 2601 Elliott Avenue, Suite 4185
 :
 : CITY: Seattle
 :
 : STATE: WA

```

1  COUNTRY: USA
2  ZIP: 98121
3
4  COMPUTER READABLE FORM:
5  MEDIUM TYPE: Diskette
6  COMPUTER: IBM Compatible
7  OPERATING SYSTEM: DOS
8  SOFTWARE: FASTCUE for Windows Version 2.0
9  CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US/08/975,316
11
12 FILING DATE:
13
14 CLASSIFICATION: 800
15
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: 08/713,000
18 FILING DATE: September 11, 1996
19
20 ATTORNEY/AGENT INFORMATION:
21
22 NAME: SLEATH, Janet
23
24 REGISTRATION NUMBER: 37,007
25
26 REFERENCE/DOCKET NUMBER: 11000/1003C1
27
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: 206-269-0565
30 TELEFAX: 206-269-0563
31
32 TELEX:
33
34 INFORMATION FOR SEQ ID NO: 9:
35
36 SEQUENCE CHARACTERISTICS:
37
38 LENGTH: 624 base pairs
39 TYPE: nucleic acid
40 STRANDEDNESS: single
41
42 TOPOLOGY: linear
43
44 US-08-975-316-9

```

Query Match	9.68;	Score 235.2;	DB 2;	Length 624;
Best Local Similarity	69.7%;	Pred. No. 3.7e-50;		
Matches 318; Conservative	0;	Mismatches 138;	Indels 0;	Gaps 0;

OY	845	TTTCGATTACGCCGAAGAGGGCTTAGCACTGTGTAACGGACACCGCGGTGGGGTCCGG	904
Db	169	TTTAACTGCAGCCCAAGGAAGACTGGCTCTGTGTACAGGGACACCGGTGGGATCCGCC	228
OY	905	ATGGCTTTCAGATGGTTCATTTGATGCTAATGTACTTCGGCTGTGTGTGCGAAGCTGTATCG	964
Db	229	GTGGCCCGGTCCTGCTCTGTGTGACGCCCAAGTCTGTGGGGTGTGTGCTGAGATCTGTCT	288
OY	965	GGCATCTTCCTGTGAGGTTATGCAAGGAAACCGGAGATTATACGATCACTTGACATCAAA	1024
Db	289	GGGCTCTTCTCGAGGCTGATGCAAGGAAACCGGAGTGTGTAATCCGTTAACCCACAG	348
OY	1025	TTTGAAACATCAACCTTGCTCAAAATCGAGCGCGCGCATATGAGATATATTTCGACGGA	1084
Db	349	TTGAAAGCACCAACCGAGGCGATGCAAGCGCGGCGGTATGATGTCTCTCTGCANCGT	408
OY	1085	AGCCATTACGTCAAGCGCGCGCAAAAGTCCACGAATGAGACCCGTTACAGAAACCAAA	1144
Db	409	AGGCACATCGTGAAGAAAGAGAGCGCGCTTCACAGAAAGACCCGTTGAGCAACCGAAA	468
OY	1145	CAAGATCGTATGCTCTCCGTACATCTTCCTCCATTTGGTGTGGATTTTAATGCAAGTATC	1204
Db	469	CAACACCGCTAGCGCTCTGGCAACATGCGCCACATGTTGGGGCTCTCCGATCCGAAGTCATC	528
OY	1205	CGATCATCAACCAAAATGATCGAGAGGAAATCAATTCGTCACACACCAACCATTTGATC	1264
Db	529	CGCGCTGTACTACTTCCTACATCGAGCGGAGATCAATTCGTCACAGCAATCCGTTATTC	588
OY	1265	GAGCTTTCGAGAAACAAAGCTTTACACCGGTGTAAAC	1300
Db	589	GATGCTTCACAGCAATGCGCTGTCTCCAGCGCGGCAAC	624

RESULT 13
US-09-211-710-9
; Sequence 9, Application US/09211710A
; Patent No. 6204434
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.

QY 1801 CTTCTGCTGTTGATGTTAAATGTTTTCCTTACATGAGAGGTTTCAGAGGTC 1860
 DB 1801 CTTCTGCTGTTGATGTTAAATGTTTTCCTTACATGAGAGGTTTCAGAGGTC 1860
 QY 1861 ATAAATATTAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 1920
 DB 1861 ATAAATATTAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 1920
 QY 1921 AACGAGACAGACAGACAGACAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 1980
 DB 1921 AACGAGACAGACAGACAGACAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 1980
 QY 1981 GAAAGTCTGTTGATGTTAAATGTTTTCCTTACATGAGAGGTTTCAGAGGTC 2040
 DB 1981 GAAAGTCTGTTGATGTTAAATGTTTTCCTTACATGAGAGGTTTCAGAGGTC 2040
 QY 2041 GTTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 2100
 DB 2041 GTTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 2100
 QY 2101 GCACTGCGACAGACAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 2160
 DB 2101 GCACTGCGACAGACAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 2160
 QY 2161 GCACTGCGACAGACAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 2220
 DB 2161 GCACTGCGACAGACAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 2220
 QY 2221 AAGTCTGTTGATGTTAAATGTTTTCCTTACATGAGAGGTTTCAGAGGTC 2280
 DB 2221 AAGTCTGTTGATGTTAAATGTTTTCCTTACATGAGAGGTTTCAGAGGTC 2280
 QY 2281 TGTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 2340
 DB 2281 TGTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 2340
 QY 2341 CATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 2400
 DB 2341 CATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 2400
 QY 2401 ATTATTTTCTGTTAAATGTTTTCCTTACATGAGAGGTTTCAGAGGTC 2460
 DB 2401 ATTATTTTCTGTTAAATGTTTTCCTTACATGAGAGGTTTCAGAGGTC 2460

RESULT 2

ID ABR50680 standard; cDNA: 2380 BP.
 XX ABR50680:
 AC 13-AUG-2002 (first entry)
 DE cDNA encoding lettuce LSPAL2 enzyme.
 KW lettuce; phenylalanine ammonia-lyase; plant wounding; plant damage;
 KM plant disease; plant browning; vegetable crop; plant; LSPAL2;
 KM gene; ss.
 XX Lactuca salvia.
 OS
 XX
 FH Key location/Qualifiers
 FT CDS 77..2218
 FT /tag a
 FT /product "LSPAL2"
 FT /transl_except- (pos:135..145, aa:Val-Gly)
 XX
 FN W0200226028-A2.
 PD 04-APR-2002.
 XX
 PF 26-SEP-2001; 2001W0-NS30192.

XX 26-SEP-2000; 2000NS-235956P
 FH
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX
 PI Salveit M, Campos R, Nonogaki H, Suslow T;
 XX WPI: 2002-435152/46.
 DB P-PDB; AAU97131.
 PT Novel polynucleotide useful in preparing plant expression constructs to
 PT modify the response of a plant cell to wounding, damage or other injury
 PT from disease causing organisms or from plant pests
 PS Claim 2; fig 7; 46pp; English.
 CC The present invention relates to the isolation of novel lettuce
 CC (lactuca salvia) phenylalanine ammonia lyase (LSPAL) enzymes, and the
 CC polynucleotide sequences encoding them. An antibody that binds to
 CC a PAL enzyme is useful for measuring the relative amount of PAL levels
 CC in a tissue. The polynucleotide sequences encoding PAL are useful in
 CC the preparation of a plant expression constructs for modifying features
 CC of the response of a plant cell to wounding, damage or other
 CC injury from disease. Disruption or down-regulation of PAL activity
 CC can be used to reduce the browning response to wounding in vegetable
 CC crops. The present sequence encodes lettuce LSPAL2 enzyme.
 CC
 SQ Sequence 2380 BP; 663 A, 548 C, 603 G; 566 T; 0 other;
 Query Match 65.7%, Score 2603.8, Lb 24, Length 2380;
 Best Local Similarity 85.5%; Pred. No. 4e-278;
 Matches 1812; Conservative 0; Mismatches 296; Indels 12; Gaps 2;
 QY 169 GATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 228
 DB 145 GATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 204
 QY 229 GAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 288
 DB 205 GAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 264
 QY 289 AGTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 348
 DB 265 TGTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 321
 QY 349 GTTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 408
 DB 322 GTTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 381
 QY 409 TAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 468
 DB 382 TAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 441
 QY 469 TAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 528
 DB 442 GAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 501
 QY 529 CAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 588
 DB 502 TAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 561
 QY 589 GAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 648
 DB 562 GAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 621
 QY 649 GAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 708
 DB 622 GAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTTTCAGAGGTC 681
 QY 709 GTTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 768
 DB 682 GTTATTTTAAATGAGAGAGTTCAGAGGTTTCCTTACATGAGAGGTTTCAGAGGTC 741

PR 06-MAY-1999: 9905-0132486-
PR 06-MAY-1999: 9905-0132487-
PR 07-MAY-1999: 9905-0132483-
PR 11-MAY-1999: 9905-0134256-
PR 14-MAY-1999: 9905-0134218-
PR 14-MAY-1999: 9905-0134219-
PR 14-MAY-1999: 9905-0134221-
PR 14-MAY-1999: 9905-0134221-
PR 14-MAY-1999: 9905-0134270-
PR 18-MAY-1999: 9905-0134768-
PR 19-MAY-1999: 9905-0134941-
PR 20-MAY-1999: 9905-0135124-
PR 21-MAY-1999: 9905-0135353-
PR 24-MAY-1999: 9905-0135629-
PR 25-MAY-1999: 9905-0136021-
PR 27-MAY-1999: 9905-0136392-
PR 28-MAY-1999: 9905-0136782-
PR 01-JUN-1999: 9905-0137222-
PR 03-JUN-1999: 9905-0137528-
PR 04-JUN-1999: 9905-0137502-
PR 07-JUN-1999: 9905-0137724-
PR 08-JUN-1999: 9905-0138094-
PR 10-JUN-1999: 9905-0138540-
PR 10-JUN-1999: 9905-0138847-
PR 14-JUN-1999: 9905-0139119-
PR 16-JUN-1999: 9905-0139452-
PR 16-JUN-1999: 9905-0139453-
PR 17-JUN-1999: 9905-0139492-
PR 18-JUN-1999: 9905-0139454-
PR 18-JUN-1999: 9905-0139455-
PR 18-JUN-1999: 9905-0139456-
PR 18-JUN-1999: 9905-0139457-
PR 18-JUN-1999: 9905-0139458-
PR 18-JUN-1999: 9905-0139459-
PR 18-JUN-1999: 9905-0139460-
PR 18-JUN-1999: 9905-0139461-
PR 18-JUN-1999: 9905-0139462-
PR 18-JUN-1999: 9905-0139463-
PR 18-JUN-1999: 9905-0139760-
PR 18-JUN-1999: 9905-0139763-
PR 21-JUN-1999: 9905-0139817-
PR 22-JUN-1999: 9905-0139899-
PR 23-JUN-1999: 9905-0140353-
PR 23-JUN-1999: 9905-0140354-
PR 24-JUN-1999: 9905-0140645-
PR 28-JUN-1999: 9905-0140823-
PR 29-JUN-1999: 9905-0140991-
PR 30-JUN-1999: 9905-0141287-
PR 01-JUL-1999: 9905-0141842-
PR 01-JUL-1999: 9905-0142154-
PR 02-JUL-1999: 9905-0142055-
PR 06-JUL-1999: 9905-0142390-
PR 08-JUL-1999: 9905-0142803-
PR 09-JUL-1999: 9905-0142920-
PR 12-JUL-1999: 9905-0142977-
PR 13-JUL-1999: 9905-0143542-
PR 14-JUL-1999: 9905-0143624-
PR 15-JUL-1999: 9905-0144085-
PR 16-JUL-1999: 9905-0144085-
PR 16-JUL-1999: 9905-0144086-
PR 19-JUL-1999: 9905-0144425-
PR 19-JUL-1999: 9905-0144431-
PR 19-JUL-1999: 9905-0144432-
PR 19-JUL-1999: 9905-0144433-
PR 19-JUL-1999: 9905-0144434-
PR 19-JUL-1999: 9905-0144435-
PR 20-JUL-1999: 9905-0144435-
PR 20-JUL-1999: 9905-0144432-
PR 20-JUL-1999: 9905-0144484-
PR 21-JUL-1999: 9905-0144614-
PR 21-JUL-1999: 9905-0145086-
PR 21-JUL-1999: 9905-0145088-
PR 22-JUL-1999: 9905-0145085-
PR 22-JUL-1999: 9905-0145087-

PR 22-JUL-1999: 9905-0145089-
PR 22-JUL-1999: 9905-0145192-
PR 23-JUL-1999: 9905-0145145-
PR 23-JUL-1999: 9905-0145218-
PR 23-JUL-1999: 9905-0145224-
PR 26-JUL-1999: 9905-0145476-
PR 27-JUL-1999: 9905-0145513-
PR 27-JUL-1999: 9905-0145518-
PR 27-JUL-1999: 9905-0145519-
PR 28-JUL-1999: 9905-0145951-
PR 02-AUG-1999: 9905-0146386-
PR 02-AUG-1999: 9905-0146388-
PR 02-AUG-1999: 9905-0146389-
PR 03-AUG-1999: 9905-0147038-
PR 04-AUG-1999: 9905-0147204-
PR 04-AUG-1999: 9905-0147302-
PR 05-AUG-1999: 9905-0147192-
PR 05-AUG-1999: 9905-0147260-
PR 06-AUG-1999: 9905-0147303-
PR 06-AUG-1999: 9905-0147416-
PR 09-AUG-1999: 9905-0147493-
PR 09-AUG-1999: 9905-0147935-
PR 10-AUG-1999: 9905-0148171-
PR 11-AUG-1999: 9905-0148319-
PR 12-AUG-1999: 9905-0148341-
PR 13-AUG-1999: 9905-0148565-
PR 13-AUG-1999: 9905-0148684-
PR 16-AUG-1999: 9905-0149368-
PR 17-AUG-1999: 9905-0149175-
PR 18-AUG-1999: 9905-0149426-
PR 20-AUG-1999: 9905-0149722-
PR 20-AUG-1999: 9905-0149723-
PR 20-AUG-1999: 9905-0149929-
PR 23-AUG-1999: 9905-0149902-
PR 23-AUG-1999: 9905-0149930-
PR 25-AUG-1999: 9905-0150566-
PR 26-AUG-1999: 9905-0150884-
PR 27-AUG-1999: 9905-0151065-
PR 27-AUG-1999: 9905-0151066-
PR 30-AUG-1999: 9905-0151080-
PR 30-AUG-1999: 9905-0151303-
PR 31-AUG-1999: 9905-0151438-
PR 01-SEP-1999: 9905-0151430-
PR 07-SEP-1999: 9905-0152363-
PR 10-SEP-1999: 9905-0153070-
PR 13-SEP-1999: 9905-0153758-
PR 15-SEP-1999: 9905-0154039-
PR 16-SEP-1999: 9905-0154039-
PR 20-SEP-1999: 9905-0154779-
PR 22-SEP-1999: 9905-0155139-
PR 23-SEP-1999: 9905-0155486-
PR 24-SEP-1999: 9905-0155486-
PR 28-SEP-1999: 9905-0156458-
PR 29-SEP-1999: 9905-0156596-
PR 04-OCT-1999: 9905-0157117-
PR 05-OCT-1999: 9905-0157753-
PR 06-OCT-1999: 9905-0157865-
PR 07-OCT-1999: 9905-0158029-
PR 08-OCT-1999: 9905-0158332-
PR 12-OCT-1999: 9905-0158469-
PR 13-OCT-1999: 9905-0159293-
PR 13-OCT-1999: 9905-0159294-
PR 13-OCT-1999: 9905-0159295-
PR 14-OCT-1999: 9905-0159329-
PR 14-OCT-1999: 9905-0159330-
PR 14-OCT-1999: 9905-0159331-
PR 14-OCT-1999: 9905-0159637-
PR 14-OCT-1999: 9905-0159638-
PR 18-OCT-1999: 9905-0159584-
PR 21-OCT-1999: 9905-0160741-
PR 21-OCT-1999: 9905-0160767-
PR 21-OCT-1999: 9905-0160768-
PR 21-OCT-1999: 9905-0160770-

PR 21-OCT-1999: 9905-0160814.
 PR 21-OCT-1999: 9905-0160815.
 PR 22-OCT-1999: 9905-0160980.
 PR 22-OCT-1999: 9905-0160981.
 PR 22-OCT-1999: 9905-0160989.
 PR 25-OCT-1999: 9905-0161404.
 PR 25-OCT-1999: 9905-0161405.
 PR 25-OCT-1999: 9905-0161406.
 PR 26-OCT-1999: 9905-0161359.
 PR 26-OCT-1999: 9905-0161360.
 PR 26-OCT-1999: 9905-0161361.
 PR 28-OCT-1999: 9905-0161920.
 PR 28-OCT-1999: 9905-0161922.
 PR 28-OCT-1999: 9905-0161993.
 PR 29-OCT-1999: 9905-0162142.

Query Match 49.8%; Score 1215.8, DB 21; Length 2400,
 Best Local Similarity 72.8%, Prd. No. 126-208;
 Matches 1584; Conservative 0; Mismatches 587; Indels 6; Gaps 1;

QY 140 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 199
 DB 148 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 207
 QY 200 GGTGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259
 DB 208 CAATGCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
 QY 260 GTGCTGAGCTGAGGAG 319
 DB 268 GTGCTGAGCTGAGGAG 321
 QY 320 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 379
 DB 322 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 381
 QY 380 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 439
 DB 382 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 441
 QY 440 GGTGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 499
 DB 442 GGTGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 501
 QY 500 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 559
 DB 502 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 561
 QY 560 GGTGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 619
 DB 562 GGTGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
 QY 620 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 679
 DB 622 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 681
 QY 680 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 739
 DB 682 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 741
 QY 740 GGTGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 799
 DB 742 GGTGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
 QY 800 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 859
 DB 802 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 861
 QY 860 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 919
 DB 862 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 921
 QY 920 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 979

DB 922 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 981
 QY 980 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1039
 DB 982 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1041
 QY 1040 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1099
 DB 1042 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1101
 QY 1100 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1159
 DB 1102 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1161
 QY 1160 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1219
 DB 1162 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1221
 QY 1220 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1279
 DB 1222 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1281
 QY 1280 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1339
 DB 1282 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1341
 QY 1340 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1399
 DB 1342 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1401
 QY 1400 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1459
 DB 1402 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1461
 QY 1460 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1519
 DB 1462 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1521
 QY 1520 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1579
 DB 1522 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1581
 QY 1580 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1639
 DB 1582 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1641
 QY 1640 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1699
 DB 1642 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1701
 QY 1700 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1759
 DB 1702 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1761
 QY 1760 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1819
 DB 1762 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1821
 QY 1820 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1879
 DB 1822 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1881
 QY 1880 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1939
 DB 1882 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1941
 QY 1940 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1999
 DB 1942 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 2001
 QY 2000 AATGAGTCTTAAATGAGTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 2059

PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150366.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	99US-0139829.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154009.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157793.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144335.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-01


```

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145413.
PR 27-JUL-1999; 99US-0145418.
PR 28-JUL-1999; 99US-0145491.
PR 28-JUL-1999; 99US-0145591.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

```

```

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157173.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 44.8%; Score 1095; DB 21; Length 2211;

Best Local Similarity 71.5%; Pred. No. 5e-187; Mismatches 580; Indels 12; Gaps 3;

```

QY 170 GATCATTAAGTGGGAGTTCAGGGAGGGGTTGACCGGAGTCACTGATGACGTG 229
DB 46 TATTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
QY 230 AAGACATGCTTGGCAGTTCAGAAACCCGTTGTAAGCTCGAGACGACCTTACA 289
DB 106 AAGCAATGTTGAAGAGATAGGAAGAGGGGTGAAGTTAGAGGTGAGACTTACG 165
QY 290 GTTTCAGGTGGCGGGATTCGCACTCTTAATGACATGACACCGTGAAGCTGACCTG 349
DB 166 ATTGTCAGTAC-----CGCGTGGCTAGAGGACGACGATCTACCGTGGACCTA 219
QY 350 TCGGAAGTCGCGAGGAGTGAAGTGAAGGCGAGTGAAGTGAAGTGAAGTGAAGTGA 409
DB 220 CGGAGAGAGCGTCTGCGGAGTGAAGGCGAGTGAAGGCGAGTGAAGTGAAGTGAAGTGA 279
QY 410 AAGCACTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 469
DB 280 GAGGAGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 339
QY 470 AAGCAAGCGGCTGCTTGAAGAGAGGCTGATTAAGTGAAGTGAAGTGAAGTGAAGTGA 529
DB 340 AAGCAAGCGGCTGCTTGAAGAGAGGCTGATTAAGTGAAGTGAAGTGAAGTGAAGTGA 399
QY 530 AATGGAAT---GGAAGCAAGCAAGTCAATGATGATGATGATGATGATGATGATGATG 586
DB 400 CCGGCGCGCGGAGACAGTCAAGACAGTCAAGACAGTCAAGACAGTCAAGACAGTCAAG 459
QY 587 GTTCAGATCAACCCCTCTCTGAGGTTTACCTCGGATGATGATGATGATGATGATGATG 646
DB 460 GTTCGTCAACCTCTCTCTCAAGGCTACTCGGATGATGATGATGATGATGATGATGATG 519

```


xx
Sequence 5291 BP; 1512 A; 1061 C; 971 G; 1747 T; 0 other;

Query Match 41.2%; Score 1005.2; DB 11; Length 5291;
Best Local Similarity 73.3%; Pred. No. 6.6e-111;
Matches 186; Conservative 0; Mismatches 468; Indels 0; Gaps 0;

501 TTGATTTTGAAGCCCGGATATATGCGCAATGCAACGGAACCAAGCCACACTTCCAC 560
3376 TCAGGTTTGTGATGCTGGGATATTTGCAATGTGTACAGATGCCAAGTCCACCTTACCC 3435
561 ATTCAAGCCACGAGCCCGCATGATGCTGCAATCAACACCTTCTGAGGCTTACCC 620
3436 AACAGTCAACTAGACGACCTATGCTTGAGAGGAACACCTTCCAGAGGACTACG 3495
621 GCATCCGATTGAGATCTTGGAGAGCATCACAGTTCCTTAACACACATCAACCCCTT 680
3496 GAATTAGATTCAATTTTGGAGGCAATCAGCAAGCTTTTCAACACACATTAACCAT 3555
681 GTTACCCCTCCGTGGAACCATCAACGCGCTCCGCTGACCTTTGCTCCATTATCATG 740
3556 GTTTCGCACTAAGGGTACATTAACGACATCTGTATCTTACTTTGTATCATATG 3615
741 CCGGCTTCTTACCGGCGCGCCCAACTGCAAAAGCCGCTGGCGCCGACCGGAGAGTCTCA 800
3616 CCGGTTTCTTAACTGATAGGCAACACTGCAAGGCTGTGTGCTCCCTGCGAGCATTTCTGA 3675
801 ATGCGGAAAAGGCTTCTTAACTGAGGATTAAGGATGATTTGAGATTAACGCA 860
3676 ATGCTAAGGAAGCCCTTGAATGAGGCAACATGCTTCTGAGCTTCTTGAATTCACCACT 3735
861 AAGAAGGCTAGCACTTGTAAAGCGGACGCGCGCTGGGCTGGGATGCGTTCATGCTTC 920
3736 AGGAAGGCTGCTGCTTGTATGATGAGAGGCTGTGTGCTGCTGCTGAGGCTATGCTTC 3795
921 TATTGATCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3855
3796 TCTTTGAAGGAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3855
981 TTATGCAAGGGAAGGCGGATTTACGATGCTGCAACAAATTTGAAGCATCAACCTG 1040
3856 TGATGCAAGGGAAGGCGGATTTACGATGCTGCAACAAATTTGAAGCATCAACCTG 3915
1041 GTCAAAATGAGGCGGCGGATGATGAGATATTTGAGGGAAGGCTTACGTCAGG 1100
3916 GCCAGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3975
1101 CGGCGCAAAAGGCTGCAAGAAATGAGCGCTTACAGAAACCAAAAGAGATGCTTATGCTC 1160
3976 CTCTTAAGAGGTTGATGAGATGATGCTTTCAGAAACCAAAAGAGATGCTTATGCTC 4035
1161 TCGGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220
4036 TTGAACTTACCCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4095
1221 TGAATGAGGGAATGCAATTCGCTCAACGACAAACCAATGATGCAAGCTTCCAGAAACA 1280
4096 CAATTGAGGGAATGCAATTCGCTCAACGACAAACCAATGATGCAAGCTTCCAGAAACA 4155
1281 AAGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340
4156 AGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4215
1341 GTCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1400
4216 GTTGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4275
1401 ATTTCACAAAGATGATGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1460
4276 ACTATATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4335
1461 ACGGCTCAAAAGGCTGCAAAATGCGCAATGCTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCT 1520

Db 4336 ATGCTTTCAAGGAGCTGAAATTTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4395
1521 CAATGCAAGTATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1580
4396 CTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4455
1581 TCGGATTAATTTGCAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGA 1640
4456 TGGGATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4515
1641 CGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
4516 CAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4575
1701 AATGCAAGTATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1760
4576 AGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4635
1761 ACGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1820
4636 ATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4695
1821 AATAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1880
4696 ACTAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4755
1881 TCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940
4756 TAAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4815
1941 CCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2000
4816 GCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4875
2001 AAGTTAAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2060
4876 AAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4935
2061 AAGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2120
4936 AGGATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2180
4996 TGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5055
2181 GCAAAAGTCAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2240
5056 GCCAAGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5115
2241 TTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5175
5116 TTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5235

RESULT 9
AAC61194
ID AAC61194 standard: DNA: 4415 BP.
AC AAC61194;
XX
AC AAC61194;
XX
12-FEB-2001 (first entry)
XX
DE Pea PEPAL DNA sequence SEQ ID 9.
XX
KW Pea; promoter; plant; prevention; pathogen; infection; ds.
XX
OS Pisum sativum.
XX
PN JP2000245463-A.
XX
PD 12-SRP-2000.
XX

Db 4219 TTCCAAATTCCTTAAGTATGATTTTAAAGAAAATTAAT-TGATCATATATATGATG 4277
OY 2301 TTTTCTTTTCTTTTAAATTAATTAATTCATTAATTCATCAAGACTTCACCTTTC 2360
Db 4278 TATTGCATATATACATAGCTGTCTCATGCTTTTAAATTAATATATATATGAAATTCATTCG 4337
OY 2361 AAGTGTGTGTATGTGTATTAATCATATATATTAATTAATTAAT 2406
Db 4338 AATTTCTTTTATGTCATGTCATGAACTGTGAATTTACTTTATATAT 4383

RESULT 10

AA046293 standard; DNA; 4415 bp.

AA046293:

09-DEC-1993 (first entry)

phenylalanine ammonia lyase gene with 5' regulatory gene.

Regulation; phenylalanine ammonia lyase; PAL; pea; ss.

Plasmid sativum.

JP05153978-A.

22-JUN-1993.

29-NOV-1991: 91JP-0317126.

29-NOV-1991: 91JP-0317126.

(TAKS) TAKASAGO PERFUMERY CO. LTD.

WPI: 1993-231495/29.

Regulator gene of specified sequence - and phenylalanine ammonia lyase gene of pea with specified sequence

Claim 2: Page 7-8; 9pp; Japanese.

The regulatory fragment of the phenylalanine ammonia lyase gene is capable of artificially accelerating or suppressing the transcription of the structural gene of a plant.

Sequence 4415 bp; 1415 A; 725 C; 793 G; 1482 T; 0 other;

Query Match 39.8%; Score 970.8; DB 14; Length 4415;

Best local Similarity 70.0%; Prod. No. 9.4e-165; Matches 1335; Conservative 0; Mismatches 567; Indels 4; Gaps 2;

OY 501 TTGATTTTGAAGCGGGAATATATGAAATGAAAGGAAACAAAGCAACATTCCTCAG 560
Db 2482 TCAGGTTTGTGAATGCTGGAATATTTGAAATGAACTGAGTCAAGCCATACATACAC 2541
OY 561 ATTACCCCAACAGCCCGCATGATGTCAGATCAACACCTCTCTCAGGCTTACTCCG 620
Db 2542 AACAGCAACAAAGGCTGAAATGATGAGAAATCAACACATTTCTTAAAGTTATTG 2601
OY 621 GCATCGATTCGATCTGGAAGCAATGACAACTGCTTAAACAAACATCACCTT 680
Db 2602 GATTTGATTTGAAATCTTGGAAGCTATTAACCAACTATTAACACACCTCACCTCAT 2661
OY 681 GTTACCCCTCGTGAACATTAATGCTGAGTGTGACCTTGTCCATTAATCATACATCG 740
Db 2662 GTTTA...CTCGGTGATCAATCAAGCTTTCGAGATTTAGTCTCTTTTATTATG 2718
OY 741 CCGGCGCTTTAAGCGCGCCCGCCCAACATCAAGCGCTTGCGCCCGGAGCAAGCTC 800
Db 2719 CTGGTTTA-TAAAGGAAATTAATTAAGGCTCATGAGACCTCTCTGCGAATTTCTTA 2778

OY 801 ATGCGAAGAAAGGCTTGTGCTGAGCGGAAATGAAAGGCTTGTGCTGAGTACAGCGGA 860
Db 2779 ATGCAAAAGAACCTTTTCATGACGCTGAAATCATATATGTTTCTTTGAATTTGCAACCA 2838
OY 861 AAGAGGCTTACACCTTTGTTAAAGCGACCCCGGCGGCTCCGGATGCTTCATGCTTC 920
Db 2839 AAGAGGCTTGTGCTTGTAAATGGAACGCTGTTGCTGCTTGTACTTCAATGTC 2898
OY 921 TATTGATGCTAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 980
Db 2899 TATTGATGCTAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2958
OY 981 TTATGCAAGGGAAGCCGAGTTTACCGATCTGACACAGAAATGAAGCATCACCTG 1040
Db 2959 TATGCAAGGGAAGCCGAGTTTACCGATCTGACACAGAAATGAAGCATCACCTG 3018
OY 1041 GTCAAAATCGAGCGCGCGGCTGATGAGTATTTTGGACGGAAGCATTTACGTCAAG 1100
Db 3019 GTCAAAATCGAGCGCGGCTGATGAGTATTTTGGACGGAAGCATTTACGTCAAG 3078
OY 1101 CGCGCGAAGAGGCTGCAAGAAATGAGACCCGTTACAGAACCAAGCATGCTGATGCTG 1160
Db 3079 CAGCTAAGAGGCTGCAAGAAATGAGACCCGTTACAGAACCAAGCATGCTGATGCTG 3138
OY 1161 TCGGTATCTCTCCGTAATGCTGAGCTGCAATGCAAGTATTCGATCATCAACAAA 1220
Db 3139 TCGGTATCTCTCCGTAATGCTGAGCTGCAATGCAAGTATTCGATCATCAACAAA 3198
OY 1221 TGAATGAGAGGAAATCAATTCGCTCAACGCAACCCATGATGAGCTTTCGAAACA 1280
Db 3199 CAATTTGAGAGGAAATCAATTCGCTCAACGCAACCCATGATGAGCTTTCGAAACA 3258
OY 1281 AAGCTTTACAGCGCTGTAATCTTCAAGAACCCCATGAGCTTTCGATGAGCAACCC 1340
Db 3259 AAGCTTTGATGCTGTAATCTTCAAGAACCCCATGAGCTTTCGATGAGCAACCC 3318
OY 1341 GTCTGCTATGCTGTAATCTTCAAGAACCCCATGAGCTTTCGATGAGCAACCC 1400
Db 3319 GTCTGCTATGCTGTAATCTTCAAGAACCCCATGAGCTTTCGATGAGCAACCC 3378
OY 1401 ATTTTACAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1460
Db 3379 ATTTTACAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3438
OY 1461 ACGGCTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1520
Db 3439 ACGGCTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3498
OY 1521 CAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1580
Db 3499 CAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3558
OY 1581 TCGGATTAATTTTACGAGAGGAAACCCGCAAGAGCTGATGATGATGATGATGATGATG 1640
Db 3559 TCGGATTAATTTTACGAGAGGAAACCCGCAAGAGCTGATGATGATGATGATGATGATG 3618
OY 1641 CGACATCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1700
Db 3619 CGACATCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3678
OY 1701 AATCGACATGAGGAAACCCGCAAGAGCTGATGATGATGATGATGATGATGATGATG 1760
Db 3679 AATCGACATGAGGAAACCCGCAAGAGCTGATGATGATGATGATGATGATGATGATG 3738
OY 1761 ACGGCGAGCTGCAAGCCGCTGAGATCTGCGAAGAAAGATGCTGCTGCTGCTGCTGCTG 1820
Db 3739 ACGGCGAGCTGCAAGCCGCTGAGATCTGCGAAGAAAGATGCTGCTGCTGCTGCTGCTG 3798
OY 1821 AATAGCTCTTGCCTTACATGACGAGCTTTCGACGCGGACATACCATTAATGCAAGAC 1880
Db 3799 AATAGCTCTTGCCTTACATGACGAGCTTTCGACGCGGACATACCATTAATGCAAGAC 3858
OY 1881 TCGGATTAATTTTACGAGAGGAAACCCGCAAGAGCTGATGATGATGATGATGATGATG 1940

OY 1518 TCCGAATCCAGTCACCAACATGTTCAAGCCCGGCAACACACATGACGTTAAT 1577
 DB 2126 TAGTAATATTCAGTACCTACCTATGTAATGTAATGTAATGTAATGTAATGTAAT 2185
 OY 1578 CTCGCGATTAATTTCCAGGAGAGAAACCCGACGACGACGACGACGACGACGACG 1637
 DB 2186 CTTTGGGCTTAATTTGGGCTTTGAAAACCTGTGAGAGCCATACAGATTAATTAAGCTTATGCT 2245
 OY 1638 GGTGCGATATCTAGTGTCTATGTCGAATTCATGATTTACGCTATTGGAAGAAC 1697
 DB 2246 CTTTCCACTATCTGTTGAGCTCTCCCAAGCTATTGATTTGAGCGATTTGAGGAGAAATTT 2305
 OY 1698 TCAAAATCGACAGTGAAGACACCGCTAAGCCAGTCCGCAAAAACGCTCCACATGCGCG 1757
 DB 2306 TCAAGAACACCTGTCAGAAATACCTGTAAGAGTTGACATGTAAGAACATTAATCTAGTGAAG 2365
 OY 1758 TCACGGCGACCTCCACCGCTGAGATTTCTGGCAGAAACATCTCTCCGCTGTTCTTATTC 1817
 DB 2366 ACAAGACAGAAACCTAACCCATTTGATTCAGTACGAGGAGAAACCTCTTAAGTGGTGATTA 2425
 OY 1818 GTCAATACCTCTGCTTATCTGTAAGATGTTGTAAGGATGTAAGGATGTAAGGATGTAAG 1877
 DB 2426 GAGAAATATGATTTTATATCATTTGATGATGATGATGATGATGATGATGATGATGATG 2485
 OY 1878 AGCTCCGACAGGTTCTGTCGACACACCGCTTAACACACGCGGCAACGAGAACACTA 1937
 DB 2486 AACTAAGAGAGTACTTTATGAGCAAGCAATACCAAGTATCAATTAATGACAAAGATGTA 2545
 OY 1938 ACACGCCATCTCCCAAAAGATGCTACTCTGCAAGAACAAATGTAAGTAACTCTGTTACCGA 1997
 DB 2546 GTTTGTTGTTTGTGGAAGATGAGCTTTTGAAGATGATGATGATGATGATGATGATGATG 2605
 OY 1998 AAGAGTTGAAGGTGTAGATGCTTATGAGATGATGATGATGATGATGATGATGATGATG 2057
 DB 2606 AGCAAGTAGAAGATGTCACCGGTAGCTTATGAGATGATGATGATGATGATGATGATGATG 2665
 OY 2058 TTTAACTTGCAGATCTACCGCTGTATAGCTTTGTAAAGGAGAGAGCTCGGACAGAGGT 2117
 DB 2666 TCAAGGAGTGAGGTATATCAGTATGTAATTTGTGAGGAGGAGGATGATGATGATGATG 2725
 OY 2118 TTTTGAACCGGACAAAGGTGAGCTGCGCGGAGAGAGCTGTCAGACAGCTGTTCCACCGGCA 2177
 DB 2726 TGTCTACCGCGGCAAAAGCTCTCTCTCCAGATGAGGAAATTTGAAAGCTTTATACAGCCA 2785
 OY 2178 TGTGCAAGAGTCAATTTATGATCGTGTGTGAGAGTGTCTGAGAGGATGAGGAGG 2237
 DB 2786 TGTGTCACGCAAAAGATTAATTTGATCCATTTGCAATGCTAGAAAGATTTGAAACGGCTTC 2845
 OY 2238 CTCCTCCATATGTTAGCAAAAGTGTGAGAGTGTGAAACGCTTTGATTTGATTTGATTTCT 2297
 DB 2846 CCATGCCAATATAGTGTTTAATGTAATTTGTTTACAAATACAGCCAGAGAAATTTT 2905
 OY 2298 GTTTTGT 2331
 DB 2906 CTCGTGTATGATTTGCTGT 2939
 RESULT 12
 AA033019
 ID AA033019 standard; cDNA; 1707 BP.
 XX
 AC AA033019;
 XX
 DT 07-MAY-1993 (first entry)
 XX
 DE Pisum sativum L. PAL gene.
 XX
 KW Phenylalanine ammonia lyase; phenyl propanoid isochlorogenic acid; plant;
 XX
 OS Pisum sativum.
 XX

PN JP04330285-A.
 PD 18-NOV-1992.
 XX
 PF 26-APR-1991; 91JP-0097697.
 XX
 PR 26-APR-1991; 91JP-0097697.
 XX
 PA (TAKS) TAKASAGO PERFUMERY CO LTD.
 DR WPI: 1993-003497/01.
 XX
 PT New phenylalanine ammonia lyase gene from *Pisum sativum* L.
 PT plasmid contg. gene, and *E. coli* transformed with plasmid, for
 PT prepn. of lyase.
 PS
 XX Disclosure: Page 8; 9pp; Japanese.
 CC The sequence is that of the phenylalanine ammonia lyase (PAL) gene
 CC from *Pisum sativum* L. The gene may be used to produce PAL.
 CC recombinantly. Plants contg. the PAL gene can contain a large amt.
 CC of phenyl propanoid isochlorogenic acid.
 CC See also AA031985.
 XX
 SQ Sequence 1707 BP; 496 A; 345 C; 377 G; 489 T; 0 other;
 Query Match 34.4%; Score 841; DB 14; Length 1707;
 Best Local Similarity 72.6%; Pred. No. 1-5e-11;
 Matches 1101; Conservative 0; Mismatches 415; Indels 1; Gaps 1;
 OY 737 ATGCGCGGCTCTTAACCG 795
 DB 1 ATGCGGCTTGTCTCACTGCGAGACCTAATCCAAAGCTGTGAGACCTCTGAGAGTAGT 60
 OY 796 CCGTAATGCGGAAAGGCGCTTGTGCGAGCGCGGAGTTAAGTGGTCTTCCGATGACA 855
 DB 61 ACTTACTGCTAAGCAAGCTTTGAAATGCGTAACTCAATTTGATGATGATGATGATGATGATG 120
 OY 856 GCGGAAGAGGAGGCTAGCACTTTTACGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 915
 DB 121 ACCCAAGAGAGGCGCTTGCCTTGTATGCGACGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 180
 OY 916 CGTTCTATTTGATGCTAATGCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 975
 DB 181 AGTTCTTTGATGCAAAATATAGTGTGTGTGTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
 OY 976 TAGGTTATGCAAGGAGGAGCGGAGTTTACGATCACTGTGACACACAAATTTGAAGATCA 1035
 DB 241 CGAAGTATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 OY 1036 CCCTGCTCAATGCAAGGCGGCGCGGATCATGAGTATTTTGGAGGAGGAGGAGGAGGAGGAGG 1095
 DB 301 CCTGCTCAATTTGAGGCGGCTGTGCGCATTTAGAGACATTTTGGATGAGAGTCTTCAT 360
 OY 1096 CAAGGCGGCGGCAAAAGGTGCGCAAGAAATGAGCCCGTTACAGAACCAACCAACATCTCTTA 1155
 DB 361 GAGGATGCTAAGAGCTGATGAAATAGATCCGTTGCAAAAGGCAAAACGAGGATGATATA 420
 OY 1156 TGTCTTGTGATCTTCTGCTTGTGATGCTTGTGATGCTTGTGATGCTTGTGATGCTTGTGATGCT 1215
 DB 421 TGCCTGCAACACTTGCACAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 480
 OY 1216 CAAATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1275
 DB 481 CAAGTCAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
 OY 1276 AAACAAGCTTTACACGCGGTAACCTTCCAGGAGAACCCGCAATCGAGTTTCCATGAGCAAA 1335
 DB 541 GAAACAAGGATTTGATGCTGATTTTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 OY 1336 CACCGGCTTGCAGATTTGCTGCAATGGAACCTGATGCTGCAATTTTCTGAGCTGGT 1395
 DB 601 CACTGCTTGCCTTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660

GenCore version 5.1.4.15_4578
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using SW model

```
Run on:      March 28, 2003, 23:41:56, Search time 67 Seconds
              (without alignments)
              1414.048 Million cell updates/sec
```

Title	US-09-96A-992A-1
Perfect score	3644
Sequence	1 MFGNHVNGVNEICIKDP, 1 FPGLEHVGSGNGBEPJLQ 711

Scoring table	BLOSUM62
Gapop 10 0	Gapext 0 5

Scattered: 908470 seeps, 133250620 residues

Parameter	Value
Total number of hits satisfying chosen parameters	908476

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing 1st 45 summaries

1. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1985.DAT *
2. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1981.DAT *
3. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1982.DAT *
4. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1983.DAT *
5. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1984.DAT *
6. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1985.DAT *
7. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1986.DAT *
8. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1987.DAT *
9. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1988.DAT *
10. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1989.DAT *
11. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1990.DAT *
12. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1991.DAT *
13. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1992.DAT *
14. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1993.DAT *
15. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1994.DAT *
16. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1995.DAT *
17. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1996.DAT *
18. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1997.DAT *
19. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1998.DAT *
20. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA1999.DAT *
21. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA2000.DAT *
22. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA2001.DAT *
23. /SID52/4/qcqd4a/a/genescq/genescsq_emb1/AA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3644	100.0	711	23	AA097130	Lettuce LSPAL1 ena
2	3294	90.4	712	23	AA097131	Lettuce LSPAL2 ena
3	3059.5	84.0	777	14	AA830077	Pisium sativum L. f
4	3054.5	83.6	725	21	AA836709	Arabidopsis thaliana
5	3044.5	83.5	709	21	AA836710	Arabidopsis thaliana
6	3041.5	83.5	698	21	AA836711	Arabidopsis thaliana
7	3004	82.4	711	21	AA654103	Arabidopsis thaliana
8	3004	82.4	717	21	AA654102	Arabidopsis thaliana
9	2951	81.0	681	21	AA654104	Arabidopsis thaliana
10	2907.5	79.8	746	21	AA653799	Arabidopsis thaliana

11	2807	77.0	699	21	AA653400	Arabidopsis thaliana
12	4548	72.7	558	21	AA653401	Arabidopsis thaliana
13	1892	51.9	423	21	AA621947	Arabidopsis thaliana
14	1799	49.4	424	21	ABD16402	Arabidopsis thaliana
15	1779	48.6	495	23	ABD78962	Lotium perenne Lp
16	1752	48.1	395	21	AA621948	Arabidopsis thaliana
17	1709	46.9	386	21	AA621949	Arabidopsis thaliana
18	1475	40.5	489	23	AB878983	Lotium perenne Lp
19	1345	36.9	412	21	AA816404	Plum radica PAL
20	1337	35.7	304	23	AAU92056	Protein encoded by K. glutinis PAL
21	1025	28.1	729	23	ABD07690	K. glutinis PAL
22	945	27.3	716	23	AAE25664	K. glutinis phenyl l-phenylalanine am
23	944	27.3	716	9	AA805313	Sequence of l-phenylalanine am
24	934.5	27.3	716	9	AA803141	Sequence of Rhodospirillum rubrum
25	944.5	27.3	716	9	AA810199	Rhodospirillum rubrum
26	944.5	27.3	716	23	AAE20654	Rhodospirillum rubrum
27	944.5	27.3	716	23	ABD07693	Rhodospirillum rubrum
28	944.5	27.3	716	23	AAE16389	K. toruloides PAL
29	932.5	27.2	716	10	AAE35783	Rhodospirillum rubrum
30	932.5	27.2	716	23	AAE20655	l-Phenylalanine am
31	931.5	27.2	716	23	AAE20654	Rhodospirillum rubrum
32	931.5	27.2	716	23	AAE20659	Rhodospirillum rubrum
33	931.5	27.2	716	23	AAE20667	K. glutinis phenyl l-phenylalanine am
34	931.5	27.2	716	23	AAE20668	K. glutinis phenyl l-phenylalanine am
35	930.5	27.2	716	23	AAE16390	Rhodospirillum rubrum
36	930.5	27.2	716	23	AAE20660	Rhodospirillum rubrum
37	930.5	27.2	716	23	AAE20662	Rhodospirillum rubrum
38	988	27.1	686	23	AAE20657	Rhodospirillum rubrum
39	987.5	27.1	716	16	AAE20682	Rhodospirillum rubrum
40	987.5	27.1	716	23	AAE20656	K. glutinis phenyl l-phenylalanine am
41	944.5	27.1	716	23	AAE20663	Rhodospirillum rubrum
42	944.5	27.1	716	23	AAE20661	Rhodospirillum rubrum
43	944.5	27.1	716	23	AAE207211	K. glutinis phenyl l-phenylalanine am
44	944.5	27.1	716	23	AAE20658	Rhodospirillum rubrum
45	971.5	26.7	713	23	ABD07692	K. rubra/mucicola

AI, IGM, INT'S

Accession	Source	Organism	Protein	Length (aa)	Weight (kDa)	PI	Ref
AA097130	AA097130 standard; protein; 711 AA.						
XX							
AC	AA097130;						
XX							
FT	13-APR-2002 (first entry)						
XX							
DE	lettuce lSPAL1 enzyme.						
XX							
KM	Lettuce; phenylalanine ammonia-lyase; plant wounding; plant damage;						
KM	plant disease; plant browning; vegetable crop; plant; lSPAL1;						
KM	enzyme.						
XX							
OS	Lactuca sativa.						
XX							
FN	W0200226028-A2.						
PD							
XX	04-APR-2002.						
XX							
PF	26-SEP-2001; 2001WO-US30192.						
XX							
PR	26-SEP-2000; 2000US-234565P.						
XX							
EA	(REGC) UNIV CALIFORNIA.						
XX							
P1	Salvicut M, Campos R, Nonogaki H, Suslow T;						
XX							
DR	WPI: 2002-43512/46.						
XX	N-PSDB: ABR50679.						
XX							
P1	Novel polynucleotide useful in preparing plant expression constructs to						
XX	modify the response of a plant cell to wounding, damage or other injury						

DB 245 GFFELQPKHGLAVNGTAVGSGMASMVLFPANVALLSFVLSALFAFVWQCKPPTDHL 304
 QY 301 IHLKIHHPGYLEAAALMEYILLGSDYVKAAGVHEMDLQKKQDKYALRTSPQWLGQTE 360
 DB 305 IHLKIHHPGYLEAAALMEYILLGSDYVKAAGVHEMDLQKKQDKYALRTSPQWLGQTE 364
 QY 361 VRSSTKMELEFINSVNNPLIDVSRKAIHGCPGCTPFGVSMNRIALIAAGLMPA 420
 DB 365 VRSSTKMELEFINSVNNPLIDVSRKAIHGCPGCTPFGVSMNRIALIAAGLMPA 424
 QY 421 QFSELYNDFYNNGLPSNLGSEFSLAYGFKGSEIAMAAYSELQFLANVINIIVSAD 480
 DB 425 QFSELYNDFYNNGLPSNLGSEFSLAYGFKGSEIAMAAYSELQFLANVINIIVSAD 484
 QY 481 IHDQNSLGLISARKFAVAVIOLKMSSTYVALVQSTPLPHEFNKSTYKKNVSVAK 540
 DB 485 IHDQNSLGLISARKFAVAVIOLKMSSTYVALVQSTPLPHEFNKSTYKKNVSVAK 544
 QY 541 KVLTMQVNGELHPSRFGCKDILKVDREYVAYIDVCSGTYPIMOKLRQVIVHAIANG 600
 DB 545 KVLTMQVNGELHPSRFGCKDILKVDREYVAYIDVCSGTYPIMOKLRQVIVHAIANG 603
 QY 601 FTEKNTNISIFOKIATFEDELKVLIPKEVGVPIAYENDTLSPNPKARPSYPIYFVR 660
 DB 604 --FENDAGTISIFOKISPEFEELKAVILPKFEGVRSAYESSTLTIPNRIKESRYPILYRFR 661
 QY 661 FELGKREFTLGEKVTSPGFEFDPVPTAMK-KQIDPLELFTLGGMN-EPPI 710
 DB 662 FELGKREFTLGEKVTSPGFEFDPVPTAMK-KQIDPLELFTLGGMN-EPPI 711

RESULT 3
 AAR30077
 ID AAR30077 standard; Protein: 777 AA.
 XX
 AC AAR10077:
 XX 07-MAY-1993 (first entry)
 XX
 DE Pisum sativum L. PAL.
 XX
 KM phenylalanine ammonia lyase; phenyl propionid isoflavonoid; plant;
 XX pea.
 XX
 OS Pisum sativum.
 XX
 PN JP04330285-A.
 XX
 PD 18-NOV-1992.
 XX
 PE 26-APR-1991: 91JP-0097697.
 XX
 PR 26-APR-1991: 91JP-0097697
 XX
 PA (TAKS) TAKASAGO PPPPUMRRY CO LTD
 XX
 DR WPI: 1993-003497/01.
 DR N-PSDB: AA031985.
 XX
 PT New phenylalanine ammonia lyase gene from Pisum sativum L. -
 PT plasmid contg. gene, and E. coli transformed with plasmid, for
 PT prepn. of lysate.
 XX
 PS Disclousure: Page 7; 9pp; Japanese.
 XX
 CC The sequence is that of phenylalanine ammonia lyase (PAL)
 CC from Pisum sativum L.. The gene encoding PAL may be used to
 CC produce PAL recombinantly. Plants contg. the PAL gene can
 CC contain a large amt. of phenyl propionid isoflavonoid.
 XX
 SQ Sequence 777 AA:
 Query Match RA 0%: Score 1059.5; DB 14; Length 777;

Host Local Similarity: 83.8%; Prod. No. 2,10-246;
 Matches 594, Conservative 57; Mismatches 51; Indels 7; Gaps 4;
 QY 3 NNNHNGVYNELGIDKPLNVAEALTSILDEKKVAVHEKPVKLGSELLIYNQVA 62
 DB 74 NAKNNMMKVNS--ADPIMGCAAEAMKGSHEVKKRVEVEKRPVVRIGLQELISQVA 140
 QY 63 GIAANDSITVYVFLSPARACVAKASSIDVMSNMKNKGTDSYCTTGCAISHPETKQGA 122
 DB 131 ATAAHDHG--KVELSESAPAVKASSIDVMSNMKNKGTDSYCTTGCAISHPETKQGA 188
 QY 123 LQKELIRFLNAGIPGNGTETSHLPASHAKAAMIVRINTLQGYSGIRFLIATIKFN 182
 DB 189 LQKELIRFLNAGIPGNGTETSHLPASHAKAAMIVRINTLQGYSGIRFLIATIKFN 248
 QY 183 NNITPCLPLRGITTAAGLVLPLSTAGLITGPNSKAVGPTGEVLNAEKFAAGVGGF 242
 DB 249 NNVTPTCL--LKGPTTASGHPVPISTAGLITGPNSKAVGPTGEVLNAEKFAAGVGGF 302
 QY 243 FHLQPKHGLAVNGTAVGSGMASMVLFPANVALLSFVLSALFAFVWQCKPPTDHLTHK 302
 DB 308 FHLQPKHGLAVNGTAVGSGMASMVLFPANVALLSFVLSALFAFVWQCKPPTDHLTHK 367
 QY 303 LKHHGQULIAAALMEYILLGSDYVKAAGVHEMDLQKKQDKYALRTSPQWLGQTE 362
 DB 368 LKHHGQULIAAALMEYILLGSDYVKAAGVHEMDLQKKQDKYALRTSPQWLGQTE 427
 QY 363 PSTKMELEFINSVNDPLIDVSRKAIHGCPGCTPFGVSMNRIALIAAGLMPA 422
 DB 428 PSTKMELEFINSVNDPLIDVSRKAIHGCPGCTPFGVSMNRIALIAAGLMPA 487
 QY 423 SELVNDYFNNGLPSNLGSEFSLAYGFKGSEIAMAAYSELQFLANVINIIVSAD 481
 DB 488 SELVNDYFNNGLPSNLGSEFSLAYGFKGSEIAMAAYSELQFLANVINIIVSAD 547
 QY 482 NQIVNSGTLISAKTAPAVVILKIMSSIVYVACOSTPPLPHEFNKSTYKKNVSVAK 541
 DB 548 NQIVNSGTLISAKTAPAVVILKIMSSIVYVACOSTPPLPHEFNKSTYKKNVSVAK 607
 QY 542 VLTMGVNEFELHPSRFGCKDILKVDREYVAYIDVCSGTYPIMOKLRQVIVHAIANG 601
 DB 608 VLTMGVNEFELHPSRFGCKDILKVDREYVAYIDVCSGTYPIMOKLRQVIVHAIANG 667
 QY 602 TREKNTNISIFOKIATFEDELKVLIPKEVGVPIAYENDTLSPNPKARPSYPIYFVR 661
 DB 668 TREKNTNISIFOKIATFEDELKVLIPKEVGVPIAYENDTLSPNPKARPSYPIYFVR 727
 QY 662 ELGRGHTLGEKVTSPGFEFDPVPTAMK-KQIDPLELFTLGGMN-EPPI 710
 DB 728 ELGRGHTLGEKVTSPGFEFDPVPTAMK-KQIDPLELFTLGGMN-EPPI 776

RESULT 4
 AAG36709
 ID AAG36709 standard; Protein: 725 AA.
 XX
 AC AAG36709:
 XX 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SHU 10 No: 45026.
 XX
 KM Protein identification: signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PR 25-FEB-2000; 2000HP-0301439.

XX
PR 25-FEB-1999; 9905-0121825-
PR 05-MAR-1999; 9905-0123180-
PR 09-MAR-1999; 9905-0123568-
PR 23-MAR-1999; 9905-0125788-
PR 25-MAR-1999; 9905-0126264-
PR 29-MAR-1999; 9905-0126785-
PR 01-APR-1999; 9905-0127462-
PR 06-APR-1999; 9905-0128234-
PR 08-APR-1999; 9905-0128714-
PR 16-APR-1999; 9905-0129845-
PR 19-APR-1999; 9905-0130077-
PR 21-APR-1999; 9905-0130449-
PR 23-APR-1999; 9905-0130510-
PR 28-APR-1999; 9905-0130891-
PR 30-APR-1999; 9905-0131449-
PR 30-APR-1999; 9905-0132048-
PR 30-APR-1999; 9905-0132407-
PR 04-MAY-1999; 9905-0132484-
PR 05-MAY-1999; 9905-0132485-
PR 06-MAY-1999; 9905-0132486-
PR 07-MAY-1999; 9905-0132487-
PR 11-MAY-1999; 9905-0132863-
PR 14-MAY-1999; 9905-0134256-
PR 14-MAY-1999; 9905-0134218-
PR 14-MAY-1999; 9905-0134219-
PR 14-MAY-1999; 9905-0134221-
PR 14-MAY-1999; 9905-0134370-
PR 18-MAY-1999; 9905-0134768-
PR 19-MAY-1999; 9905-0134941-
PR 20-MAY-1999; 9905-0135124-
PR 21-MAY-1999; 9905-0135353-
PR 24-MAY-1999; 9905-0135629-
PR 25-MAY-1999; 9905-0136021-
PR 27-MAY-1999; 9905-0136392-
PR 28-MAY-1999; 9905-0136782-
PR 01-JUN-1999; 9905-0137222-
PR 03-JUN-1999; 9905-0137528-
PR 04-JUN-1999; 9905-0137502-
PR 07-JUN-1999; 9905-0137724-
PR 08-JUN-1999; 9905-0138094-
PR 10-JUN-1999; 9905-0138540-
PR 10-JUN-1999; 9905-0138847-
PR 14-JUN-1999; 9905-0139119-
PR 16-JUN-1999; 9905-0139452-
PR 16-JUN-1999; 9905-0139453-
PR 17-JUN-1999; 9905-0139492-
PR 18-JUN-1999; 9905-0139454-
PR 18-JUN-1999; 9905-0139455-
PR 18-JUN-1999; 9905-0139456-
PR 18-JUN-1999; 9905-0139457-
PR 18-JUN-1999; 9905-0139458-
PR 18-JUN-1999; 9905-0139459-
PR 18-JUN-1999; 9905-0139460-
PR 18-JUN-1999; 9905-0139461-
PR 18-JUN-1999; 9905-0139462-
PR 18-JUN-1999; 9905-0139463-
PR 18-JUN-1999; 9905-0139470-
PR 18-JUN-1999; 9905-0139763-
PR 21-JUN-1999; 9905-0139817-
PR 22-JUN-1999; 9905-0139899-
PR 23-JUN-1999; 9905-0140353-
PR 23-JUN-1999; 9905-0140354-
PR 24-JUN-1999; 9905-0140695-
PR 28-JUN-1999; 9905-0140823-
PR 29-JUN-1999; 9905-0140991-
PR 30-JUN-1999; 9905-0141287-
PR 01-JUL-1999; 9905-0141842-
PR 01-JUL-1999; 9905-0142154-
PR 02-JUL-1999; 9905-0142055-
PR 06-JUL-1999; 9905-0142390-
PR 08-JUL-1999; 9905-0142803-
PR 09-JUL-1999; 9905-0142920-

PR 12-JUL-1999; 9905-0142977-
PR 13-JUL-1999; 9905-0143342-
PR 14-JUL-1999; 9905-0143624-
PR 15-JUL-1999; 9905-0144005-
PR 16-JUL-1999; 9905-0144085-
PR 16-JUL-1999; 9905-0144086-
PR 19-JUL-1999; 9905-0144325-
PR 19-JUL-1999; 9905-0144331-
PR 19-JUL-1999; 9905-0144332-
PR 19-JUL-1999; 9905-0144333-
PR 19-JUL-1999; 9905-0144334-
PR 19-JUL-1999; 9905-0144335-
PR 20-JUL-1999; 9905-0144352-
PR 20-JUL-1999; 9905-0144632-
PR 21-JUL-1999; 9905-0144814-
PR 21-JUL-1999; 9905-0145086-
PR 21-JUL-1999; 9905-0145088-
PR 22-JUL-1999; 9905-0145085-
PR 22-JUL-1999; 9905-0145087-
PR 22-JUL-1999; 9905-0145089-
PR 23-JUL-1999; 9905-0145192-
PR 23-JUL-1999; 9905-0145145-
PR 23-JUL-1999; 9905-0145218-
PR 23-JUL-1999; 9905-0145224-
PR 26-JUL-1999; 9905-0145276-
PR 27-JUL-1999; 9905-0145913-
PR 27-JUL-1999; 9905-0145918-
PR 27-JUL-1999; 9905-0145919-
PR 28-JUL-1999; 9905-0145951-
PR 02-AUG-1999; 9905-0146386-
PR 02-AUG-1999; 9905-0146388-
PR 02-AUG-1999; 9905-0146389-
PR 03-AUG-1999; 9905-0147038-
PR 04-AUG-1999; 9905-0147204-
PR 04-AUG-1999; 9905-0147302-
PR 05-AUG-1999; 9905-0147192-
PR 05-AUG-1999; 9905-0147260-
PR 06-AUG-1999; 9905-0147303-
PR 06-AUG-1999; 9905-0147416-
PR 09-AUG-1999; 9905-0147493-
PR 09-AUG-1999; 9905-0147935-
PR 10-AUG-1999; 9905-0148171-
PR 11-AUG-1999; 9905-0148319-
PR 12-AUG-1999; 9905-0148341-
PR 13-AUG-1999; 9905-0148565-
PR 13-AUG-1999; 9905-0148684-
PR 16-AUG-1999; 9905-0149368-
PR 17-AUG-1999; 9905-0149175-
PR 18-AUG-1999; 9905-0149426-
PR 20-AUG-1999; 9905-0149722-
PR 20-AUG-1999; 9905-0149723-
PR 20-AUG-1999; 9905-0149929-
PR 23-AUG-1999; 9905-0149902-
PR 23-AUG-1999; 9905-0149930-
PR 25-AUG-1999; 9905-0150566-
PR 26-AUG-1999; 9905-0150884-
PR 27-AUG-1999; 9905-0151065-
PR 27-AUG-1999; 9905-0151086-
PR 30-AUG-1999; 9905-0151303-
PR 31-AUG-1999; 9905-0151438-
PR 01-SEP-1999; 9905-0151930-
PR 07-SEP-1999; 9905-0152363-
PR 10-SEP-1999; 9905-0153070-
PR 13-SEP-1999; 9905-0153758-
PR 15-SEP-1999; 9905-0154018-
PR 16-SEP-1999; 9905-0154039-
PR 20-SEP-1999; 9905-0154779-
PR 22-SEP-1999; 9905-0155139-
PR 23-SEP-1999; 9905-0155486-
PR 24-SEP-1999; 9905-0155659-
PR 28-SEP-1999; 9905-0156458-

PR 29-SEP-1999: 990S-0156596.
 PR 04-OCT-1999: 990S-0157117.
 PR 05-OCT-1999: 990S-0157753.
 PR 06-OCT-1999: 990S-0157865.
 PR 07-OCT-1999: 990S-0158029.
 PR 08-OCT-1999: 990S-0158232.
 PR 12-OCT-1999: 990S-0158369.
 PR 13-OCT-1999: 990S-0159294.
 PR 13-OCT-1999: 990S-0159294.
 PR 13-OCT-1999: 990S-0159294.
 PR 14-OCT-1999: 990S-0159329.
 PR 14-OCT-1999: 990S-0159330.
 PR 14-OCT-1999: 990S-0159331.
 PR 14-OCT-1999: 990S-0159637.
 PR 14-OCT-1999: 990S-0159638.
 PR 18-OCT-1999: 990S-0159584.
 PR 21-OCT-1999: 990S-0160741.
 PR 21-OCT-1999: 990S-0160747.
 PR 21-OCT-1999: 990S-0160768.
 PR 21-OCT-1999: 990S-0160770.
 PR 21-OCT-1999: 990S-0160815.
 PR 22-OCT-1999: 990S-0160980.
 PR 22-OCT-1999: 990S-0160981.
 PR 22-OCT-1999: 990S-0160989.
 PR 25-OCT-1999: 990S-0161404.
 PR 25-OCT-1999: 990S-0161405.
 PR 25-OCT-1999: 990S-0161406.
 PR 26-OCT-1999: 990S-0161359.
 PR 26-OCT-1999: 990S-0161360.
 PR 26-OCT-1999: 990S-0161361.
 PR 28-OCT-1999: 990S-0161920.
 PR 28-OCT-1999: 990S-0161992.
 PR 28-OCT-1999: 990S-0161993.
 PR 29-OCT-1999: 990S-0162142.

Query Match 83.6%; Score 3045.5; DB 21: Length 725;
 Best Local Similarity 81.4%; Pred. No. 2,562,35;
 Matches 589, Conservative 52, Mismatches 66, Indels 17, Gaps 3;

OY 3 NCHNHN---GVNNGIC-----IKDPLNMGVAAEFALITGSHIDREYKKMAEERKP 47
 DB 4 NCHNHN---GVNNGIC-----IKDPLNMGVAAEFALITGSHIDREYKKMAEERKP 43
 OY 48 VVAKIGCELTIVSQAAGIAAANISDITVKEVLSIAARAGYKASSDWMVSMNKGUSGYVTF 107
 DB 64 VVAKIGCELTIVSQAAGIAAANISDITVKEVLSIAARAGYKASSDWMVSMNKGUSGYVTF 107
 OY 108 GRSATSHPTKQYALDYKELTFELNAGTFGNTTETSHITLPHSATPAMTVPINTLLQYYS 167
 DB 122 GRSATSHPTKQYALDYKELTFELNAGTFGNTTETSHITLPHSATPAMTVPINTLLQYYS 167
 OY 168 GTRFELFAITFKLNNNTTPCLYRGCTTASGDIVPLSYIAGLITGRNSNAAVPTGCVL 227
 DB 182 GTRFELFAITFKLNNNTTPCLYRGCTTASGDIVPLSYIAGLITGRNSNAAVPTGCVL 227
 OY 228 NAKKAAAGVHGFELQPRKQALVNGTAVGSCMASMVLFEANVLALESFVLSATIPAE 287
 DB 242 NAKKAAAGVHGFELQPRKQALVNGTAVGSCMASMVLFEANVLALESFVLSATIPAE 287
 OY 288 VMOGKREFTDHLTHKIKHHPQOIEAAMTMEYILDSGDVYKAAQVHENDPLQKQJRYA 347
 DB 302 VMOGKREFTDHLTHKIKHHPQOIEAAMTMEYILDSGDVYKAAQVHENDPLQKQJRYA 347
 OY 348 LRTSPQMLAQHOLVIRSSSTKMEEREINSVNDNPLIDVSRKNALKQGNFOGTPIGVSMINT 407
 DB 362 LRTSPQMLAQHOLVIRSSSTKMEEREINSVNDNPLIDVSRKNALKQGNFOGTPIGVSMINT 407
 OY 408 KIALIAAIGKLMIAOFSHIVNDYVNNGLJNSI.SGGRNNSLDYGFQKQIAMAASYSSEHIOPL 467
 DB 422 KIALIAAIGKLMIAOFSHIVNDYVNNGLJNSI.SGGRNNSLDYGFQKQIAMAASYSSEHIOPL 467
 OY 468 ANFVTHHVSADQHNDVNSIGLISARKTAEAVDILKLMSSITYLVALGOSIDLHREEM 527

DB 482 ANPVTSHVGSADQHNDVNSIGLISARKTAEAVDILKLMSSITYLVALGOSIDLHREEM 541
 OY 528 KSTVKNVSOVAKKVLITGVNGEELAPSRFCEKDLARVDREYVAVIDVCSGLYPLMOK 587
 DB 542 KSTVKNVSOVAKKVLITGVNGEELAPSRFCEKDLARVDREYVAVIDVCSGLYPLMOK 587
 OY 588 LQVLYVDAALNGEETKNTNITFOKIALTEFEELKVLIPKEVEGVIAVENDLSTIPNKT 647
 DB 602 LQVLYVDAALNGEETKNTNITFOKIALTEFEELKVLIPKEVEGVIAVENDLSTIPNKT 647
 OY 648 KCPSPYLPYRFPFELSGPPTTGEKVTSGDEEDPVFPAFKQOIDPLELUGMNGEP 707
 DB 662 KCPSPYLPYRFPFELSGPPTTGEKVTSGDEEDPVFPAFKQOIDPLELUGMNGEP 707
 OY 708 LPIG 711
 DB 722 LPIG 725
 RESULT 5
 ID AAG36710 standard; protein; 709 AA.
 AC AAG36710;
 XX 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 45027.
 XX
 KW Hybridisation; signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 OS Arabidopsis thaliana.
 XX
 PN EP103405-A2.
 XX
 PU 06-SEP-2000.
 XX
 PR 25-FEB-2000: 2000EP-0301439.
 XX
 PR 25-FEB-1999: 990S-0121825.
 PR 05-MAR-1999: 990S-0123180.
 PR 09-MAR-1999: 990S-0123548.
 PR 23-MAR-1999: 990S-0125788.
 PR 25-MAR-1999: 990S-0126264.
 PR 29-MAR-1999: 990S-0126785.
 PR 01-APR-1999: 990S-0127462.
 PR 06-APR-1999: 990S-0128234.
 PR 08-APR-1999: 990S-0128714.
 PR 16-APR-1999: 990S-0129845.
 PR 19-APR-1999: 990S-0130077.
 PR 21-APR-1999: 990S-0130449.
 PR 23-APR-1999: 990S-0130510.
 PR 23-APR-1999: 990S-0130891.
 PR 28-APR-1999: 990S-0131449.
 PR 30-APR-1999: 990S-0132044.
 PR 04-MAY-1999: 990S-0132407.
 PR 05-MAY-1999: 990S-0132484.
 PR 06-MAY-1999: 990S-0132485.
 PR 06-MAY-1999: 990S-0132486.
 PR 07-MAY-1999: 990S-0132487.
 PR 11-MAY-1999: 990S-0134256.
 PR 14-MAY-1999: 990S-0134219.
 PR 14-MAY-1999: 990S-0134221.
 PR 14-MAY-1999: 990S-0134270.
 PR 18-MAY-1999: 990S-0134768.
 PR 19-MAY-1999: 990S-0134941.
 PR 20-MAY-1999: 990S-0135124.
 PR 21-MAY-1999: 990S-0135353.

PK	24-MAY-1999	9905-013565.2
PK	25-MAY-1999	9905-013606.21
PP	27-MAY-1999	9905-013636.92
PP	28-MAY-1999	9905-013676.82
PP	01-JUN-1999	9905-013722.22
PP	03-JUN-1999	9905-013752.28
PP	04-JUN-1999	9905-013767.02
PP	07-JUN-1999	9905-013772.24
PP	08-JUN-1999	9905-013809.94
PP	10-JUN-1999	9905-013854.40
PP	14-JUN-1999	9905-013884.77
PP	14-JUN-1999	9905-013911.19
PP	16-JUN-1999	9905-013945.52
PP	16-JUN-1999	9905-013946.53
PP	18-JUN-1999	9905-013945.57
PP	18-JUN-1999	9905-013945.58
PP	18-JUN-1999	9905-013945.59
PP	18-JUN-1999	9905-013946.60
PP	18-JUN-1999	9905-013946.61
PP	18-JUN-1999	9905-013946.62
PP	18-JUN-1999	9905-013946.63
PP	18-JUN-1999	9905-013957.50
PP	18-JUN-1999	9905-013976.53
PP	21-JUN-1999	9905-013981.77
PP	22-JUN-1999	9905-013986.99
PP	23-JUN-1999	9905-014035.93
PP	23-JUN-1999	9905-014035.94
PP	24-JUN-1999	9905-014068.93
PP	28-JUN-1999	9905-014086.25
PP	29-JUN-1999	9905-014099.91
PP	30-JUN-1999	9905-014128.87
PP	01-JUL-1999	9905-014184.42
PP	02-JUL-1999	9905-014205.54
PP	08-JUL-1999	9905-014239.00
PP	08-JUL-1999	9905-014239.01
PP	08-JUL-1999	9905-014239.02
PP	08-JUL-1999	9905-014239.27
PP	12-JUL-1999	9905-014287.77
PP	13-JUL-1999	9905-014354.42
PP	14-JUL-1999	9905-014362.24
PP	15-JUL-1999	9905-014400.05
PP	15-JUL-1999	9905-014408.65
PP	15-JUL-1999	9905-014408.66
PP	15-JUL-1999	9905-014434.25
PP	19-JUL-1999	9905-014434.31
PP	19-JUL-1999	9905-014433.32
PP	19-JUL-1999	9905-014433.33
PP	19-JUL-1999	9905-014433.34
PP	19-JUL-1999	9905-014434.34
PP	19-JUL-1999	9905-014435.32
PP	20-JUL-1999	9905-014456.32
PP	20-JUL-1999	9905-014468.84
PP	22-JUL-1999	9905-014508.89
PP	22-JUL-1999	9905-014519.92
PP	23-JUL-1999	9905-014514.45
PP	23-JUL-1999	9905-014521.84
PP	23-JUL-1999	9905-014522.84
PP	26-JUL-1999	9905-014557.76
PP	27-JUL-1999	9905-014593.13
PP	27-JUL-1999	9905-014593.18
PP	28-JUL-1999	9905-014595.19
PP	02-AUG-1999	9905-014636.88
PP	02-AUG-1999	9905-014636.88

FR	02-AUG-1999	9905-0146389
FR	03-AUG-1999	9905-0147030
FR	04-AUG-1999	9905-0147204
FR	04-AUG-1999	9905-0147302
FR	05-AUG-1999	9905-0147192
FR	05-AUG-1999	9905-0147760
FR	05-AUG-1999	9905-0147303
FR	06-AUG-1999	9905-0147416
FR	09-AUG-1999	9905-0147493
FR	09-AUG-1999	9905-0147935
FR	10-AUG-1999	9905-0148171
FR	11-AUG-1999	9905-0148341
FR	12-AUG-1999	9905-0148545
FR	13-AUG-1999	9905-0148684
FR	16-AUG-1999	9905-0149366
FR	17-AUG-1999	9905-0149175
FR	18-AUG-1999	9905-0149426
FR	20-AUG-1999	9905-0149722
FR	20-AUG-1999	9905-0149723
FR	20-AUG-1999	9905-0149929
FR	22-AUG-1999	9905-0149902
FR	23-AUG-1999	9905-0149930
FR	25-AUG-1999	9905-0150566
FR	26-AUG-1999	9905-0150884
FR	27-AUG-1999	9905-0151065
FR	27-AUG-1999	9905-0151066
FR	30-AUG-1999	9905-0151003
FR	31-AUG-1999	9905-0151103
FR	01-SEP-1999	9905-0151338
FR	01-SEP-1999	9905-0151519
FR	07-SEP-1999	9905-0152263
FR	10-SEP-1999	9905-0153070
FR	13-SEP-1999	9905-0153758
FR	15-SEP-1999	9905-0154039
FR	16-SEP-1999	9905-0154039
FR	20-SEP-1999	9905-0154779
FR	22-SEP-1999	9905-0155131
FR	24-SEP-1999	9905-0155486
FR	26-SEP-1999	9905-0156458
FR	26-SEP-1999	9905-0156566
FR	04-OCT-1999	9905-0157117
FR	05-OCT-1999	9905-0157753
FR	06-OCT-1999	9905-0157865
FR	07-OCT-1999	9905-0158239
FR	08-OCT-1999	9905-0158232
FR	12-OCT-1999	9905-0158393
FR	13-OCT-1999	9905-0158293
FR	13-OCT-1999	9905-0159294
FR	13-OCT-1999	9905-0159295
FR	14-OCT-1999	9905-0159329
FR	14-OCT-1999	9905-0159331
FR	14-OCT-1999	9905-0159637
FR	14-OCT-1999	9905-0159638
FR	16-OCT-1999	9905-0159584
FR	21-OCT-1999	9905-0160741
FR	21-OCT-1999	9905-0160767
FR	22-OCT-1999	9905-0160981
FR	25-OCT-1999	9905-0161008
FR	25-OCT-1999	9905-0161405
FR	25-OCT-1999	9905-0161405
FR	26-OCT-1999	9905-0161406
FR	26-OCT-1999	9905-0161360
FR	26-OCT-1999	9905-0161361
FR	26-OCT-1999	9905-0161360

PR 28-OCT-1999: 9905-0161992.
PR 28-OCT-1999: 9905-0161993.
PR 29-OCT-1999: 9905-0161994.

Query Match 83.5%; Score 3044.5; PR 21; Length 109;
Host Local Similarity 83.1%; Pred. No. 2,90-235;
Matches 585, Conservatively 52, Mismatches 62, Indels 5, Gaps 2.

QY 8 NNVNDELCTIDPLNMGVAEALTSGLIDFVKKMAEFKKPVKKLGGFTITVQVAGIAAA 67
DB 11 NNVNDELCTIDPLNMGVAEALTSGLIDFVKKMAEFKKPVKKLGGFTITVQVAGIAAA 67
QY 68 NDSDFVAVELSPFAKGVKASDPMWESMKNKIDSYGVTCGASHKKTOKGALQKEL 127
DB 68 GNS--YKVELSETARAGVANNSSDMWESMKNKIDSYGVTCGASHKKTOKGALQKEL 125
QY 128 IRELNAGIFPNKGTETSTPLHSTRAAMIVRINTLLQYSSGTFPELEATIKELNNITP 187
DB 126 IRELNAGIFPNKGTETSTPLHSTRAAMIVRINTLLQYSSGTFPELEATIKELNNITP 185
QY 188 CTPLRFTITASGLVPELVSTATLLTGRPNKAVGPTGELNKKAAVAGGCGFEELQP 247
DB 186 STPLRFTITASGLVPELVSTATLLTGRPNKAVGPTGELNKKAAVAGGCGFEELQP 245
QY 248 KEGIALAVNGIAVSGMASMVLIDAVNLAISHVLSAFAVWQKREFTDHLTKIKLHHP 307
DB 246 KEGIALAVNGIAVSGMASMVLIDAVNLAISHVLSAFAVWQKREFTDHLTKIKLHHP 305
QY 308 GUTFAAAIMETITLGSFYKAAKVIEMKPLKPKDRAKALISYQWLPGUJEVIRKSTK 367
DB 306 GUTFAAAIMETITLGSFYKAAKVIEMKPLKPKDRAKALISYQWLPGUJEVIRKSTK 365
QY 368 MFEELINSVNDPLIDVSPKALHFGNFQITPLGSMNTPLAIAAGIKLMAQFSELYN 427
DB 366 STEELINSVNDPLIDVSPKALHFGNFQITPLGSMNTPLAIAAGIKLMAQFSELYN 425
QY 428 DFTNNGLPSMLSGCKNSLYLQCKKGEIAMAISGSLQPLIAPVINVHVSARHONQVNS 487
DB 426 DFTNNGLPSMLSGCKNSLYLQCKKGEIAMAISGSLQPLIAPVINVHVSARHONQVNS 485
QY 488 LGLISAKTAEADITLKLMSSTYLVATGSDPLPHLEFNKSTVKNKTVSQVAKKYLTCGV 547
DB 486 LGLISAKTAEADITLKLMSSTYLVATGSDPLPHLEFNKSTVKNKTVSQVAKKYLTCGV 545
QY 548 NGELHPSRFCEKDLRVNDEVFAVTDVGSCTGYPLMKELPQVIVPAALNGETEKNTN 607
DB 546 NGELHPSRFCEKDLRVNDEVFAVTDVGSCTGYPLMKELPQVIVPAALNGETEKNTN 605
QY 608 TSTIFOKIATPEEELKVLKPKVGCVRIAVFNPTLSIPNRIKACRSVPVIRPVREELQGP 667
DB 606 TSTIFOKIATPEEELKVLKPKVGCVRIAVFNPTLSIPNRIKACRSVPVIRPVREELQGP 665
QY 668 LTGEKVTSPGEETRVFTAMKQGITPLIEVLSGNNLEPLPIC 711
DB 666 LTGEKVTSPGEETRVFTAMKQGITPLIEVLSGNNLEPLPIC 709

RESULT 6
AAG36711 standard, Protein, 698 AA.
AC AAG36711;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SPQ ID NO: 45028.
XX
XX
XX protein identification: signal transduction pathway, metabolic pathway;
XX hybridisation assay, genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX

PN HP1033405-A2.
XX 06-SEP-2000
FE 25-FEB-2003; 2000EP-0301439.
XX 25-FEB-1999: 9905-0121825.
PR 05-MAR-1999: 9905-012180.
PR 09-MAR-1999: 9905-0123548.
PR 23-MAR-1999: 9905-0125788.
PR 26-MAR-1999: 9905-0126264.
PR 29-MAR-1999: 9905-0126785.
PR 01-APR-1999: 9905-0127462.
PR 06-APR-1999: 9905-0128234.
PR 08-APR-1999: 9905-0128714.
PR 16-APR-1999: 9905-0129845.
PR 19-APR-1999: 9905-0130077.
PR 21-APR-1999: 9905-0130449.
PR 23-APR-1999: 9905-0130510.
PR 28-APR-1999: 9905-0130891.
PR 30-APR-1999: 9905-0131449.
PR 30-APR-1999: 9905-0132048.
PR 30-APR-1999: 9905-0132407.
PR 04-MAY-1999: 9905-0132484.
PR 05-MAY-1999: 9905-0132485.
PR 05-MAY-1999: 9905-0132486.
PR 06-MAY-1999: 9905-0132487.
PR 07-MAY-1999: 9905-0132863.
PR 11-MAY-1999: 9905-0134256.
PR 14-MAY-1999: 9905-0134218.
PR 14-MAY-1999: 9905-0134219.
PR 14-MAY-1999: 9905-0134221.
PR 14-MAY-1999: 9905-0134370.
PR 18-MAY-1999: 9905-0134768.
PR 19-MAY-1999: 9905-0134941.
PR 20-MAY-1999: 9905-0135124.
PR 21-MAY-1999: 9905-0135353.
PR 24-MAY-1999: 9905-0135629.
PR 25-MAY-1999: 9905-0136021.
PR 27-MAY-1999: 9905-0136782.
PR 28-MAY-1999: 9905-0136782.
PR 01-JUN-1999: 9905-0137422.
PR 03-JUN-1999: 9905-0137528.
PR 04-JUN-1999: 9905-0137502.
PR 07-JUN-1999: 9905-0137724.
PR 08-JUN-1999: 9905-0138094.
PR 10-JUN-1999: 9905-0138540.
PR 10-JUN-1999: 9905-0138847.
PR 14-JUN-1999: 9905-0139119.
PR 16-JUN-1999: 9905-0139452.
PR 16-JUN-1999: 9905-0139453.
PR 17-JUN-1999: 9905-0139492.
PR 18-JUN-1999: 9905-0139454.
PR 18-JUN-1999: 9905-0139455.
PR 18-JUN-1999: 9905-0139456.
PR 18-JUN-1999: 9905-0139457.
PR 18-JUN-1999: 9905-0139458.
PR 18-JUN-1999: 9905-0139459.
PR 18-JUN-1999: 9905-0139460.
PR 18-JUN-1999: 9905-0139461.
PR 18-JUN-1999: 9905-0139462.
PR 18-JUN-1999: 9905-0139463.
PR 18-JUN-1999: 9905-0139464.
PR 18-JUN-1999: 9905-0139465.
PR 21-JUN-1999: 9905-0139817.
PR 23-JUN-1999: 9905-0139899.
PR 23-JUN-1999: 9905-0140353.
PR 23-JUN-1999: 9905-0140354.
PR 24-JUN-1999: 9905-0140695.
PR 24-JUN-1999: 9905-0140823.
PR 24-JUN-1999: 9905-0140991.
PR 30-JUN-1999: 9905-0141287.
PR 01-JUN-1999: 9905-0141842.

```

PK 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142300.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143277.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145132.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149476.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

```

```

PK 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155446.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159658.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      83.58; Score 3041.5; DB 21; Length 698;
Best Local Similarity 81.28; Fred. No. 4.9e-235;
Matches 584; Conservative 52; Mismatches 61; Indels 5; Caps 2;

QY 10 VNNECICDPPLNMGVAAEALTGSHLDEVKKMAEERKPVYKLGGETLTVSOVAGIAAND 69
DB 2 VIN--AHDPLNMGAAAFQMKGSHLDEVKKMAEERKPVYKLGGETLTVSOVAGIAISTGN 58
QY 70 SDYKVELSEARACVAKASSDVMWESMNNKCTDSYGVITGCGATLSHKRIKQGCALQKELLR 129
DB 59 S--YKVELSEFARAGVASSDVMWESMNNKCTDSYGVITGCGATLSHKRIKQGCALQKELLR 116
QY 130 FLNNGIFGNGFTSHTLPHSATRAAMIYRIINTLLOGSGIRFELTALTKFLNNNTPL 189
DB 117 FLNNGIFGNGFTSHTLPHSATRAAMIYRIINTLLOGSGIRFELTALTKFLNNNTPL 176
QY 190 PLRGITIASGDLVPELSTAGLITGRPNKAVGPTGEVYLNKAKFAAGYGGFFELQPKRE 249
DB 177 PLRGITIASGDLVPELSTAGLITGRPNKAVGPTGEVYLNKAKFAAGYGGFFELQPKRE 236
QY 250 GLIAYNGIAVSGSMAWVLFUANYLALLSVLSAIFAEVWCKPFRIDILIRKLAKHPQV 309
DB 237 GLIAYNGIAVSGSMAWVLFUANYLALLSVLSAIFAEVWCKPFRIDILIRKLAKHPQV 296
QY 310 IEAAAMEYIIDGSDYKAAOKVHEMDLQPKQODRYALKTSQWIGPOLFYVRSSTKM 369
DB 297 IEAAAMEYIIDGSDYKAAOKVHEMDLQPKQODRYALKTSQWIGPOLFYVRSSTKM 356
QY 370 EPEINSVNDPLDIVSRNKALHGGNFGTPIGVSMNTPLAIAIGKLMFAQFSELVNDF 429
DB 357 EPEINSVNDPLDIVSRNKALHGGNFGTPIGVSMNTPLAIAIGKLMFAQFSELVNDF 416

```


PR	0.2-AUG-1999	9905-0146388
PR	0.2-AUG-1999	9905-0146389
PR	03-AUG-1999	9905-0147030
PR	04-AUG-1999	9905-0147704
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148341
PR	12-AUG-1999	9905-0148565
PR	13-AUG-1999	9905-0148684
PR	16-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149176
PR	18-AUG-1999	9905-0149122
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149929
PR	22-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150366
PR	26-AUG-1999	9905-0150884
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	07-SEP-1999	9905-0152363
PR	10-SEP-1999	9905-0153700
PR	11-SEP-1999	9905-0153758
PR	15-SEP-1999	9905-0154013
PR	16-SEP-1999	9905-0154039
PR	20-SEP-1999	9905-0155179
PR	22-SEP-1999	9905-0155133
PR	23-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0155559
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156177
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	06-OCT-1999	9905-0157645
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158332
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	13-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159430
PR	14-OCT-1999	9905-0159531
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	18-OCT-1999	9905-0159584
PR	18-OCT-1999	9905-0159680
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160767
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160770
PR	21-OCT-1999	9905-0160815
PR	22-OCT-1999	9905-0160815
PR	22-OCT-1999	9905-0160980
PR	22-OCT-1999	9905-0160981
PR	22-OCT-1999	9905-0160983
PR	25-OCT-1999	9905-0161404
PR	25-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	26-OCT-1999	9905-0161359
PR	26-OCT-1999	9905-0161360
PR	26-OCT-1999	9905-0161361

PK	28-OCT-1999;	99US-0161920
PK	28-OCT-1999;	99US-0161992
PR	28-OCT-1999;	99US-0161993
PR	29-OCT-1999;	99US-0162142

Query Match	82.48;	Score 3004;	DB 21;	Length 711;
Best Local Similarity	81.58;	Pred. No. 5.1e-232;		
Matches 577;	Conservative 52;	Mismatches 77;	Indels 2;	Gaps 1;

QY 4 GNHNVGVEELCIKPELMMGVAAELTGSILDEYKKVVAEERKPVYKLGGETLTYSQVAG 63
 Db 6 GERTKKAIVITIKTILADPLMMNGLAADQMGSHLDEYKKVVEYKKRPVYVNGGETLTIGVAA 65
 QY 64 IAAANDSVYKVELSEAPACVYKASSPMVMSMKNKGTDSYGVTTGFGATSHRRTKQCCAL 123
 Db 66 ISTRVGS--VKVELLETSTRAGYKASSDPMVMSMKNKGTDSYGVTTGFGATSHRRTKNGAL 123
 QY 124 UKELIRLVNAGIFGNGETSHLTPSATRAMIVRIINTLLOGSGIRPELTATIKPELNN 183
 Db 124 QTELIRLVNAGIFGNGETSHLTPSATRAMIVRIINTLLOGSGIRPELTATISLNN 183
 QY 184 NITPCELPREITITASGDVLPYSTIAGLTIGRPNSKAVGPGEVLAENKAFAAAGVEGFE 243
 Db 184 NISPLPLRGTITASGDVLPYSTIAGLTIGRPNSKATCPDGESLTAKAEFEKAGISTGFE 243
 QY 244 ELDPKESLALVNGTAGVSGMSMNVLPDANVALISFLSAIFAEVQCKPETHLTHTKL 303
 Db 244 DLQPKESLALVNTAGVSGMSMNVLPEDANVALISLFAEVNSGKPEETHLTHTKL 303
 QY 304 KHHGQGLEAAALIMEYTLDSQDVYKAAQVHBMPLDQPKQDRYALRTSPQWLGQIEYR 363
 Db 304 KHHGQGLEAAALIMEHTLDOSSYMKLAQVHBMPLDQPKQDRALRTSPQWLGQIEYR 363
 QY 364 SSTKMEIDREINSYNDNPFLIDVSRKALHGNFGCTPIGVSMDNTRLAIAIGKLMAQFS 423
 Db 364 QATKSIEDREINSYNDNPFLIDVSRKALHGNFGCTPIGVSMDNTRLAIAIGKLMAQFS 423
 QY 424 ELVNDPFPNNLPLPNNLSGGHNPRLDYGFCGRTIAMAQVCSFLOPLANPNTNHVQSAEOHNO 483
 Db 424 ELVNDPFPNNLPLPNNLTASSNPBLDYGFCGRTIAMAQVCSFLOPLANPNTSHVQSAEOHNO 483
 QY 484 DVNSLGLISARKIAEAVDILIKLMSSTYVVAACQSIDLRHLEEMKSTVKNVTSQYAKKVL 543
 Db 484 DVNSLGLISSRKISEAVDILIKLMSSTPLVIGICQAVDLRHLLEENLRQTVKNTVTSQYAKKVL 543
 QY 544 TMYNGSLHSPRCCEKDLRYVDPEYFVAYLIDVYCSGTYPMLKRLQVYLVDAALNGETE 603
 Db 544 TTGTINGSLHSPRCCEKDLRYVDPEYFVAYVDDCSATYPLMLRQLQVYLVDAALNSGETE 603
 QY 604 KNTFTSLFQRIATFEELIKVLLPKEVEGVRILAYENDPLSLPNRIKQNSYPLYRFRVEEL 663
 Db 604 KNAATSTFQKIGTAFBEELIKAVLPKREYLAARAAYNNGTAPLPNNIKCKRSTPLYRFRVEEL 663
 QY 664 GRGHTIGEKVTSFGRHDPVPTAMCKGOIIDLPLECLGNGNGEPRLPIC 711
 Db 664 GTKLITGEKVVSPGEFEFDKVFAMCEGLIDPLMDCLKEMNGCAPIPIC 711
 RESULT 8
 AAG54102
 ID AAG54102 standard; Protein: 717 AA.
 AC AAG54102;
 XX
 XX 18-OCT-2000 (first entry)
 XX
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 68948.
 XX
 KW Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay: genetic mapping; gene expression control; promoter
 KW termination sequence.
 XX
 XX Arabidopsis thaliana.

XX HP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PP 25-FEB-2000; 2000EP-0301434
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130444.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134376.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139464.
PR 18-JUN-1999; 990S-0139465.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140645.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.

PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142590.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0146388.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.

```

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

```

Query Match 82.48; Score 3004; Len 21; Length 717;
 Best local similarity 81.58; pred. No. 5.2e-232;
 Matches 577; Conservative 52; Mismatches 77; Indels 1;

```

QY 4 GNHNVGVNELCIRDPINMGVAEALTGSHLDEYKKVAEFRRPVYKLGSETLTVSQVAG 63
DB 12 GKTIVAVATTKTLDPLNMGIAADOMKSHLDEYKKVVEYRRPVNIAGETLTIGQVAA 71
QY 64 IAAANDSTTVVVELSEAFARVAKSSDMWESNMKKITLVS;VTTT;FAT;SHPTK;S;AL 123
DB 72 ISTVGS--KVELAETSRVAVKSSDMWESNMKKITD;SVTTG;FAT;SHPTK;S;AL 129
QY 124 QKEILKFNAGIPQNGEISHILPHSAIKRAMIKINILAGYSGILFELLEAFKFLNN 183
DB 130 QTELIFLNATIFPNTKETCHITLPOSATRAAMLVFVNTLLQ;SSGIFELLEAITSLNH 189
QY 184 NITPDLPLRGITITASGLVPLSTYATGLTTPPNKSAVGPTEGVLMKAKFAAAGVEGGEF 243
DB 190 NISPSILPKGITITASGLVPLSTYATGLTTPPNKSAVGPTEGVLMKAKFAAAGVEGGEF 249
QY 244 ELOPEEGALVNGTAVGSGMNSWLPEDANVLLISEVLSITFEVNGQKPEFDHLLIHL 303
DB 250 DLQPEEGALVNGTAVGSGMNSWLPEDANVLLISEVLSITFEVNGQKPEFDHLLIHL 309
QY 304 KHHQDQIFAAALMEYILDGSDYVAAQKVHMDPLQKPKQDKAKIKISVQMLQVLEVR 363
DB 310 KHHQDQIFAAALMEYILDGSDYVAAQKVHMDPLQKPKQDKAKIKISVQMLQVLEVR 369
QY 364 SSTKIEKEINSVNDPLIDVSRKKALHGNFQGTPIGVSMONTRLALAAIGKLMRAQFS 423
DB 370 QATKSIEKEINSVNDPLIDVSRKKALHGNFQGTPIGVSMONTRLALAAIGKLMRAQFS 429

```

```

QY 424 ELVNDPVNNGLPNLSCGRNPSLDYQFKGGEIAMASYCSHLQPLAMPVNHVQSAVQHNQ 483
DB 430 ELVNDPVNNGLPNLSSNPSLDYQFKGGEIAMASYCSHLQPLAMPVNHVQSAVQHNQ 489
QY 484 DVNSIGLISARKTAEAVDILKMSSTYVALCOSIDLRLHLENNKSTVANTVSQAKVYL 543
DB 490 DVNSIGLISARKTAEAVDILKMSSTYVALCOSIDLRLHLENNKSTVANTVSQAKVYL 549
QY 544 TMSVNGELHPSKFEKELLVVEFEYFAYIDV;S;STTPLM;K;LV;V;V;H;A;L;N;K;E;T;E 603
DB 550 TTGNGELHPSKFEKELLVVEFEYFAYIDV;S;STTPLM;K;LV;V;V;H;A;L;N;K;E;T;E 609
QY 604 KNTNTSIFOKIATFEEDLKYLLPEKEVEGVRAVENDTLSIPNTRKASRSPLYRFVREEL 663
DB 610 KNAVTSIFOKIATFEEDLKYLLPEKEVEGVRAVENDTLSIPNTRKASRSPLYRFVREEL 669
QY 664 GRGELTGKVTSPGEERDRVFTAMCKGQITIDPLFCLGCMNGRPDIPIC 711
DB 670 GTRKLTGKVTSPGEERDRVFTAMCKGQITIDPLFCLGCMNGRPDIPIC 717

RESULT 9
AAG54104
ID AAG54104 standard; Protein; 681 AA.
AC AAG54104;
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment Sbj ID No: 68950.
DE
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EPI033405-A2.
FN
XX
XX 06 SEP-2000.
CE
XX
XX 25-FEB-2000; 2000RP-0301439.
IT
XX
XX 25-FEB-1999; 99US-0121825.
ER
XX 05-MAR-1999; 99US-0123180.
PR
XX 09-MAR-1999; 99US-0123548.
PR
XX 23-MAR-1999; 99US-0125788.
PR
XX 25-MAR-1999; 99US-0126264.
PR
XX 29-MAR-1999; 99US-0126785.
PR
XX 01-APR-1999; 99US-0127462.
PR
XX 06-APR-1999; 99US-0128234.
PR
XX 08-APR-1999; 99US-0128714.
PR
XX 16-APR-1999; 99US-0129845.
PR
XX 19-APR-1999; 99US-0130077.
PR
XX 21-APR-1999; 99US-0130449.
PR
XX 23-APR-1999; 99US-0130510.
PR
XX 23-APR-1999; 99US-0130891.
PR
XX 28-APR-1999; 99US-0131449.
PR
XX 30-APR-1999; 99US-0132048.
PR
XX 30-APR-1999; 99US-0132407.
PR
XX 04-MAY-1999; 99US-0132484.
PR
XX 05-MAY-1999; 99US-0132485.
PR
XX 06-MAY-1999; 99US-0132486.
PR
XX 06-MAY-1999; 99US-0132487.
PR
XX 07-MAY-1999; 99US-0132863.
PR
XX 11-MAY-1999; 99US-0134256.
PR
XX 14-MAY-1999; 99US-0134219.
PR
XX 14-MAY-1999; 99US-0134221.
PR
XX 14-MAY-1999; 99US-0134370.
PR
XX 18-MAY-1999; 99US-0134768.
PR
XX 19-MAY-1999; 99US-0134941.

```

PR 20-MAY-1999: 99US-0137124
PR 21-MAY-1999: 99US-0135353
PR 24-MAY-1999: 99US-0135629
PR 25-MAY-1999: 99US-0136021
PR 28-MAY-1999: 99US-0136392
PR 28-MAY-1999: 99US-0136782
PR 01-JUN-1999: 99US-0137222
PR 03-JUN-1999: 99US-0137528
PR 04-JUN-1999: 99US-0137502
PR 07-JUN-1999: 99US-0137724
PR 08-JUN-1999: 99US-0138094
PR 10-JUN-1999: 99US-0138540
PR 10-JUN-1999: 99US-0138847
PR 14-JUN-1999: 99US-0139419
PR 16-JUN-1999: 99US-0139452
PR 16-JUN-1999: 99US-0139453
PR 17-JUN-1999: 99US-0139492
PR 18-JUN-1999: 99US-0139454
PR 18-JUN-1999: 99US-0139455
PR 18-JUN-1999: 99US-0139456
PR 18-JUN-1999: 99US-0139457
PR 18-JUN-1999: 99US-0139458
PR 18-JUN-1999: 99US-0139459
PR 18-JUN-1999: 99US-0139460
PR 18-JUN-1999: 99US-0139461
PR 18-JUN-1999: 99US-0139462
PR 18-JUN-1999: 99US-0139463
PR 18-JUN-1999: 99US-0139750
PR 18-JUN-1999: 99US-0139763
PR 21-JUN-1999: 99US-0139817
PR 22-JUN-1999: 99US-0139899
PR 23-JUN-1999: 99US-0140353
PR 23-JUN-1999: 99US-0140354
PR 24-JUN-1999: 99US-0140695
PR 28-JUN-1999: 99US-0140823
PR 28-JUN-1999: 99US-0140991
PR 30-JUN-1999: 99US-0141287
PR 01-JUL-1999: 99US-0141842
PR 01-JUL-1999: 99US-0142154
PR 02-JUL-1999: 99US-0142055
PR 06-JUL-1999: 99US-0142390
PR 08-JUL-1999: 99US-0142803
PR 09-JUL-1999: 99US-0142920
PR 12-JUL-1999: 99US-0142977
PR 13-JUL-1999: 99US-0143542
PR 14-JUL-1999: 99US-0143624
PR 15-JUL-1999: 99US-0144005
PR 16-JUL-1999: 99US-0144085
PR 16-JUL-1999: 99US-0144086
PR 19-JUL-1999: 99US-0144325
PR 19-JUL-1999: 99US-0144331
PR 19-JUL-1999: 99US-0144332
PR 19-JUL-1999: 99US-0144333
PR 19-JUL-1999: 99US-0144334
PR 20-JUL-1999: 99US-0144335
PR 20-JUL-1999: 99US-0144352
PR 20-JUL-1999: 99US-0144632
PR 20-JUL-1999: 99US-0144884
PR 21-JUL-1999: 99US-0144814
PR 21-JUL-1999: 99US-0145086
PR 21-JUL-1999: 99US-0145088
PR 22-JUL-1999: 99US-0145087
PR 22-JUL-1999: 99US-0145087
PR 22-JUL-1999: 99US-0145089
PR 22-JUL-1999: 99US-0145192
PR 23-JUL-1999: 99US-0145145
PR 23-JUL-1999: 99US-0145218
PR 23-JUL-1999: 99US-0145224
PR 26-JUL-1999: 99US-0145276
PR 26-JUL-1999: 99US-0145293
PR 27-JUL-1999: 99US-0145418
PR 27-JUL-1999: 99US-0145919
PR 28-JUL-1999: 99US-0145951
PR 02-AUG-1999: 99US-0145486
PR 02-AUG-1999: 99US-0146388
PR 02-AUG-1999: 99US-0146389
PR 03-AUG-1999: 99US-0147038
PR 04-AUG-1999: 99US-0147204
PR 04-AUG-1999: 99US-0147204
PR 05-AUG-1999: 99US-0147192
PR 05-AUG-1999: 99US-0147260
PR 06-AUG-1999: 99US-0147303
PR 06-AUG-1999: 99US-0147416
PR 09-AUG-1999: 99US-0147493
PR 09-AUG-1999: 99US-0147935
PR 10-AUG-1999: 99US-0148171
PR 11-AUG-1999: 99US-0148319
PR 12-AUG-1999: 99US-0148341
PR 13-AUG-1999: 99US-0148565
PR 13-AUG-1999: 99US-0148684
PR 15-AUG-1999: 99US-0149368
PR 17-AUG-1999: 99US-0149475
PR 18-AUG-1999: 99US-0149426
PR 20-AUG-1999: 99US-0149722
PR 20-AUG-1999: 99US-0149723
PR 20-AUG-1999: 99US-0149929
PR 23-AUG-1999: 99US-0149902
PR 23-AUG-1999: 99US-0149930
PR 25-AUG-1999: 99US-0150566
PR 26-AUG-1999: 99US-0150884
PR 27-AUG-1999: 99US-0151065
PR 27-AUG-1999: 99US-0151066
PR 27-AUG-1999: 99US-0151066
PR 30-AUG-1999: 99US-0151080
PR 31-AUG-1999: 99US-0151303
PR 01-SEP-1999: 99US-0151338
PR 01-SEP-1999: 99US-0151930
PR 07-SEP-1999: 99US-0152363
PR 10-SEP-1999: 99US-0153070
PR 14-SEP-1999: 99US-0153758
PR 15-SEP-1999: 99US-0154018
PR 16-SEP-1999: 99US-0154039
PR 20-SEP-1999: 99US-0154779
PR 22-SEP-1999: 99US-0155139
PR 23-SEP-1999: 99US-0155486
PR 23-SEP-1999: 99US-0155659
PR 24-SEP-1999: 99US-0156458
PR 24-SEP-1999: 99US-0156458
PR 24-SEP-1999: 99US-0156458
PR 04-OCT-1999: 99US-0157117
PR 05-OCT-1999: 99US-0157753
PR 06-OCT-1999: 99US-0157865
PR 07-OCT-1999: 99US-0158029
PR 07-OCT-1999: 99US-0158432
PR 12-OCT-1999: 99US-0158369
PR 13-OCT-1999: 99US-0159293
PR 13-OCT-1999: 99US-0159294
PR 13-OCT-1999: 99US-0159295
PR 14-OCT-1999: 99US-0159329
PR 14-OCT-1999: 99US-0159330
PR 14-OCT-1999: 99US-0159331
PR 14-OCT-1999: 99US-0159637
PR 14-OCT-1999: 99US-0159638
PR 18-OCT-1999: 99US-0159584
PR 21-OCT-1999: 99US-0160741
PR 21-OCT-1999: 99US-0160767
PR 21-OCT-1999: 99US-0160770
PR 21-OCT-1999: 99US-0160770
PR 21-OCT-1999: 99US-0160814
PR 21-OCT-1999: 99US-0160815
PR 22-OCT-1999: 99US-0160980
PR 22-OCT-1999: 99US-0160981
PR 22-OCT-1999: 99US-0160984
PR 25-OCT-1999: 99US-0161404
PR 25-OCT-1999: 99US-0161405
PR 25-OCT-1999: 99US-0161406
PR 26-OCT-1999: 99US-0161359
PR 26-OCT-1999: 99US-0161360

PR 26-OCT-1999; 9905-0161361.
 PR 28-OCT-1999; 9905-0161920.
 PR 28-OCT-1999; 9905-0161932.
 PR 28-OCT-1999; 9905-0161993.
 PR 29-OCT-1999; 9905-0162142.

Query Match 81.0%; Score 2931, DB 21, Length 681,
 Best Local Similarity 83.0%; Pred. No. 8,6e-228;
 Matches 567, Conservative 49, Mismatches 75, Gaps 1.

UY 29 LIGSHLDEKKKVAAPKPKVKKLDELITVSQVAGIAANISDVLKVELSAARAKKAS 88
 DB 1 MGSFLDEKKKVEEYRPPVNLSEETLTIGVAAITV3SS--VKVELETSHASVKAS 58
 UY 89 SIWVESNNKQIISYGVITGFAISHRETKQJCAICKELHPEINAGIEGNTSESHLPH 148
 DB 59 SIWVESNNKQIDSYGVITGFAISHRETKQJCAICKELHPEINAGIEGNTSESHLPH 118
 UY 149 SATRAAMIVRINILLYGSGIRFEILEAITKELNNNITPCLPLKGTITANSGLVPLSYIA 208
 DB 119 SATRAAMIVRINILLYGSGIRFEILEAITKELNNNITPCLPLKGTITANSGLVPLSYIA 178
 UY 209 GILTRPNSKATGCEVLENAKFAAAGVCECPPELOPKGICAIYNGTAVGSSMASMV. 268
 DB 179 GILTRPNSKATGCEVLENAKFAAAGVCECPPELOPKGICAIYNGTAVGSSMASMV. 238
 UY 269 PDANVALTLEVLSTAFVAVNCKPEPTDHLPHKIKHHIGQIFAAALMPYLNDSDVKA 328
 DB 239 FANVQVAVLAVLSAIFVAVNCKPEPTDHLPHKIKHHIGQIFAAALMPYLNDSDVKA 268
 UY 329 AKVHEHDEFLUKKQKUKALKTSYQWLPULFEVIFSSIKMLBEFINVNNPILIDVSRNK 368
 DB 299 AKVHEHDEFLUKKQKUKALKTSYQWLPULFEVIFSSIKMLBEFINVNNPILIDVSRNK 358
 UY 389 ALHGSNPGITPIGVSMUNIKLAIATIGKLMFAQPSLVNDPYNNGLSNLSGCGNPSLDY 448
 DB 359 ALHGSNPGITPIGVSMUNIKLAIATIGKLMFAQPSLVNDPYNNGLSNLSGCGNPSLDY 418
 UY 449 GFKGSEIAMASYCSLQFLANPVINHOVSARHONQVNSLQILSAKAVARAVPLIKIMSS 508
 DB 419 GFKGSEIAMASYCSLQFLANPVINHOVSARHONQVNSLQILSAKAVARAVPLIKIMSS 478
 UY 509 TYLVALGOSIDLRHEENKSTVNTVSQVAKKYLINGVNGELHPSPECKDLIRVADRE 568
 DB 479 TYLVALGOSIDLRHEENKSTVNTVSQVAKKYLINGVNGELHPSPECKDLIRVADRE 538
 UY 569 VYFAYIDVCSQTYPLMKQLKQVILVDHALNNGELKNTNTSIFOKIATPEPEIKVILPKKE 628
 DB 539 VYFAYIDVCSQTYPLMKQLKQVILVDHALNNGELKNTNTSIFOKIATPEPEIKVILPKKE 598
 UY 629 VAGVRIATENDLISIPNKIKACRSYPLIRFVREHJQGLIQKQVISPGEEDKVFIAAC 688
 DB 599 VFAAARAAGNGTAPIPNKIKACRSYPLIRFVREHJQGLIQKQVISPGEEDKVFIAAC 658
 UY 689 KGOIIDPLELLECGVNSSEPLPIC 711
 DB 659 KGOIIDPLELLECGVNSSEPLPIC 681

RESULT 10
 AAG53399
 ID AAG53399 standard; Protein: 736 AA.

AC AAG53399;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID No: n7482.

XX protein identification; signal transduction pathway; metabolic pathway;
 KW hydrolisation assay; genetic mapping; gene expression control; promoter,
 KM termination sequence.
 XX

OS Arabidopsis thaliana.

XX EPIJ33405-A2.

XX 06-SEP-2000.

XX 25 FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 9905-011945.

XX 05-MAR-1999; 9905-0124180.

XX 09-MAR-1999; 9905-0123548.

XX 23-MAR-1999; 9905-0125788.

XX 25-MAR-1999; 9905-0126264.

XX 29-MAR-1999; 9905-0126785.

XX 01-APR-1999; 9905-0127462.

XX 06-APR-1999; 9905-0128234.

XX 08-APR-1999; 9905-0128714.

XX 16-APR-1999; 9905-0129845.

XX 19-APR-1999; 9905-0130077.

XX 21-APR-1999; 9905-0130449.

XX 23-APR-1999; 9905-0130510.

XX 28-APR-1999; 9905-0130891.

XX 30-APR-1999; 9905-0132048.

XX 04-MAY-1999; 9905-0132407.

XX 05-MAY-1999; 9905-0132484.

XX 06-MAY-1999; 9905-0132485.

XX 06-MAY-1999; 9905-0132486.

XX 06-MAY-1999; 9905-0132487.

XX 07-MAY-1999; 9905-0132863.

XX 11-MAY-1999; 9905-0134256.

XX 14-MAY-1999; 9905-0134218.

XX 14-MAY-1999; 9905-0134219.

XX 14-MAY-1999; 9905-0134221.

XX 14-MAY-1999; 9905-0134370.

XX 18-MAY-1999; 9905-0134768.

XX 19-MAY-1999; 9905-0134941.

XX 20-MAY-1999; 9905-0135124.

XX 21-MAY-1999; 9905-0135353.

XX 24-MAY-1999; 9905-0135629.

XX 25-MAY-1999; 9905-0136021.

XX 27-MAY-1999; 9905-0136392.

XX 28-MAY-1999; 9905-0136782.

XX 01-JUN-1999; 9905-0137222.

XX 03-JUN-1999; 9905-0137528.

XX 04-JUN-1999; 9905-0137502.

XX 07-JUN-1999; 9905-0137724.

XX 08-JUN-1999; 9905-0138094.

XX 10-JUN-1999; 9905-0138540.

XX 10-JUN-1999; 9905-0138847.

XX 14-JUN-1999; 9905-0139119.

XX 16-JUN-1999; 9905-0139452.

XX 16-JUN-1999; 9905-0139453.

XX 17-JUN-1999; 9905-0139492.

XX 18-JUN-1999; 9905-0139454.

XX 18-JUN-1999; 9905-0139455.

XX 18-JUN-1999; 9905-0139456.

XX 18-JUN-1999; 9905-0139457.

XX 18-JUN-1999; 9905-0139458.

XX 18-JUN-1999; 9905-0139459.

XX 18-JUN-1999; 9905-0139460.

XX 18-JUN-1999; 9905-0139461.

XX 18-JUN-1999; 9905-0139462.

XX 18-JUN-1999; 9905-0139463.

XX 18-JUN-1999; 9905-0139750.

XX 18-JUN-1999; 9905-0139763.

XX 21-JUN-1999; 9905-0139817.

XX 22-JUN-1999; 9905-0139809.

XX 23-JUN-1999; 9905-0140353.

XX 23-JUN-1999; 9905-0140354.

XX 24-JUN-1999; 9905-0140695.

XX 28-JUN-1999; 9905-0140823.

XX 29-JUN-1999; 9905-0140991.

```

PR 30-JUN-1999: 990S-0141287.
PR 01-JUL-1999: 990S-0141842.
PR 01-JUL-1999: 990S-0142154.
PR 02-JUL-1999: 990S-0142364.
PR 06-JUL-1999: 990S-0142390.
PR 08-JUL-1999: 990S-0142803.
PR 09-JUL-1999: 990S-0142920.
PR 12-JUL-1999: 990S-0142977.
PR 13-JUL-1999: 990S-0143542.
PR 14-JUL-1999: 990S-0143624.
PR 15-JUL-1999: 990S-0144005.
PR 16-JUL-1999: 990S-0144046.
PR 19-JUL-1999: 990S-0144325.
PR 19-JUL-1999: 990S-0144341.
PR 19-JUL-1999: 990S-0144332.
PR 19-JUL-1999: 990S-0144334.
PR 19-JUL-1999: 990S-0144334.
PR 20-JUL-1999: 990S-0144335.
PR 20-JUL-1999: 990S-0144352.
PR 20-JUL-1999: 990S-0144632.
PR 21-JUL-1999: 990S-0144814.
PR 21-JUL-1999: 990S-0145086.
PR 22-JUL-1999: 990S-0145085.
PR 22-JUL-1999: 990S-0145087.
PR 22-JUL-1999: 990S-0145089.
PR 22-JUL-1999: 990S-0145192.
PR 23-JUL-1999: 990S-0145145.
PR 23-JUL-1999: 990S-0145218.
PR 23-JUL-1999: 990S-0145224.
PR 26-JUL-1999: 990S-0145276.
PR 27-JUL-1999: 990S-0145918.
PR 27-JUL-1999: 990S-0145919.
PR 28-JUL-1999: 990S-0145951.
PR 02-AUG-1999: 990S-0146386.
PR 02-AUG-1999: 990S-0146388.
PR 03-AUG-1999: 990S-0147038.
PR 04-AUG-1999: 990S-0147204.
PR 04-AUG-1999: 990S-0147302.
PR 05-AUG-1999: 990S-0147192.
PR 05-AUG-1999: 990S-0147260.
PR 06-AUG-1999: 990S-0147303.
PR 06-AUG-1999: 990S-0147416.
PR 09-AUG-1999: 990S-0147493.
PR 09-AUG-1999: 990S-0147935.
PR 10-AUG-1999: 990S-0148171.
PR 11-AUG-1999: 990S-0148319.
PR 12-AUG-1999: 990S-0148341.
PR 13-AUG-1999: 990S-0148565.
PR 13-AUG-1999: 990S-0148684.
PR 16-AUG-1999: 990S-0149388.
PR 17-AUG-1999: 990S-0149175.
PR 18-AUG-1999: 990S-0149426.
PR 20-AUG-1999: 990S-0149722.
PR 20-AUG-1999: 990S-0149723.
PR 20-AUG-1999: 990S-0149929.
PR 23-AUG-1999: 990S-0149902.
PR 23-AUG-1999: 990S-0149930.
PR 25-AUG-1999: 990S-0149930.
PR 26-AUG-1999: 990S-0150084.
PR 27-AUG-1999: 990S-0151065.
PR 27-AUG-1999: 990S-0151066.
PR 27-AUG-1999: 990S-0151080.
PR 30-AUG-1999: 990S-0151303.
PR 31-AUG-1999: 990S-0151438.
PR 01-SEP-1999: 990S-0151930.
PR 07-SEP-1999: 990S-0152363.
PR 10-SEP-1999: 990S-0153070.
PR 13-SEP-1999: 990S-0153758.

```

```

PR 15-SEP-1999: 990S-0154018.
PR 16-SEP-1999: 990S-0154039.
PR 20-SEP-1999: 990S-0154779.
PR 22-SEP-1999: 990S-0155139.
PR 23-SEP-1999: 990S-0155486.
PR 24-SEP-1999: 990S-0155659.
PR 28-SEP-1999: 990S-0156458.
PR 29-SEP-1999: 990S-0156596.
PR 04-OCT-1999: 990S-0157117.
PR 05-OCT-1999: 990S-0157753.
PR 06-OCT-1999: 990S-0157865.
PR 07-OCT-1999: 990S-0158029.
PR 08-OCT-1999: 990S-0158232.
PR 12-OCT-1999: 990S-0158369.
PR 13-OCT-1999: 990S-0158369.
PR 13-OCT-1999: 990S-0159294.
PR 13-OCT-1999: 990S-0159295.
PR 14-OCT-1999: 990S-0159329.
PR 14-OCT-1999: 990S-0159330.
PR 14-OCT-1999: 990S-0159331.
PR 14-OCT-1999: 990S-0159437.
PR 14-OCT-1999: 990S-0159584.
PR 18-OCT-1999: 990S-0159584.
PR 21-OCT-1999: 990S-0160741.
PR 21-OCT-1999: 990S-0160767.
PR 21-OCT-1999: 990S-0160768.
PR 21-OCT-1999: 990S-0160770.
PR 21-OCT-1999: 990S-0160814.
PR 21-OCT-1999: 990S-0160815.
PR 22-OCT-1999: 990S-0160980.
PR 22-OCT-1999: 990S-0160981.
PR 22-OCT-1999: 990S-0160981.
PR 25-OCT-1999: 990S-0161404.
PR 25-OCT-1999: 990S-0161405.
PR 25-OCT-1999: 990S-0161406.
PR 26-OCT-1999: 990S-0161359.
PR 26-OCT-1999: 990S-0161360.
PR 26-OCT-1999: 990S-0161361.
PR 28-OCT-1999: 990S-0161920.
PR 28-OCT-1999: 990S-0161992.
PR 28-OCT-1999: 990S-0161993.
PR 29-OCT-1999: 990S-0162142.

```

Query Match 79.8%, Score 2907.5, Dh 21, Length 736,

Best Local Similarity 80.7%, Fred. No. 424,

Matches 568, Conservative 50, Mismatches 77, Indels 9, Gaps 4,

```

QY 5 NHVGVNRLCTIKDPLNMCVAAEALVGSILDEYVKKMAEPRPVVKLOGPILVSOVAGI 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8 NHTTAVSG----DPLNMTATFAALKGSHLDYVKKHVEYKFAVKLOGPILVGOVAAV 62

QY 65 AANSDTYKVELSEFAAVGVKASSSEMYMESNNKKTDSYGVTTGFGATSHRETQVATLG 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 63 ARGGGSGT--VRLAFARAGVAKSSSEMYMESNNKKTDSYGVTTGFGATSHRETQVATLG 120

QY 125 KDLIRFLNAGIFGNGI-LVSHLIPPSAIPAAIVRINILDSYSGIRPELLEATIKPLNN 183
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 NELIRFLNAGITGFANQDSITLTFKTFEAMLVAVRNLDSYSGIRPELLEATIKPLNN 180

QY 184 NITPLDPLDKGTTASGDIYPLISYIAGLLIGRPNKAVGPTDCHVNAKKAFAACVGGGPF 243
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 EITPLDPLDKGTTASGDIYPLISYIAGLLIGRPNKAVGPTDCHVNAKKAFAACVGGGPF 243

QY 244 ELDPREGALVNGTAVGSMASNVLEFADNVLLSEVLSAIFAEVMGKPEFTDHTHKL 303
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 240 ELDPREGALVNGTAVGSMASNVLEFADNVLLSEVLSAIFAEVMGKPEFTDHTHKL 299

QY 304 KHHPGQIFAAAIMFYLDGSDYVKAQVYHEMDPLQKTRQDRAVAKLSPWLGQDIEVIR 363
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 300 KHHPGQIFAAAIMHEHLLDSSSYKAEVLIHEMDPLQKTRQDRAVAKLSPWLGQDIEVIR 359

QY 364 SSTKMIHREINSVNPNPLDVSNNKALHGOVNGGPPICVSMQDNRILAAIGKIMFQOPS 423
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Dh 360 AATKMEHETNSVNNPDLIVSEKKAALH33NPHTTIPGVAMNSKIALASICKIMAPQS 419
Qy 424 ELVDEFTNNGIPELNLSSGRNPSLUDYGFKGGEIMASTGSEFLPLAFVTHVSAQJHQ 483
Dh 420 ELVDEFTNNGIPELNLSSGRNPSLUDYGFKGGEIMASTGSEFLPLAFVTHVSAQJHQ 479
Qy 484 DVNSIGLISAKRTAEADYDIKIMSSYLVALCOSIDIRHLEEMKSTVKNVSQVAKKVI 543
Dh 480 DVNSIGLISAKRTAEADYDIKIMSSYLVALCOSIDIRHLEEMKSTVKNVSQVAKKVI 539
Qy 544 TMGNGELHPSRFECEKDLKRVNDREYFAVIDVCSSTPYLMKLVQVLDHALNGETE 603
Dh 540 TMGNGELHPSRFECEKDLKRVNDREYFAVIDVCSSTPYLMKLVQVLDHALNGETE 599
Qy 604 KNTNISIFOKIATFEELKVLLEPREVEGRIVAYENDPLSTPNRIFKPSYPLTFVFEEL 663
Dh 600 ANSATSVHRTKGAFAEAEKLLLPREVEGRIVAYENDPLSTPNRIFKPSYPLTFVFEEL 659
Qy 664 GRCPITGCKVTSPEGEEDRVFTAMCKQIIDPILECLGNGEHP 707
Dh 660 NTELTGENVRSPGEEDRVFTAMCKQIIDPILECLGNGEHP 703

RESULT 11
AAG53400
ID AAG53400 standard; Protein; 699 AA.
XX AAG53400:
XX
XX 18-OCT-2000 (first entry)
XX
Dh Arabidopsis thaliana protein fragment Shu IIb No: 67983.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EPI033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134708.

PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 24-MAY-1999; 9905-0135353.
PR 25-MAY-1999; 9905-0135629.
PR 27-MAY-1999; 9905-0136021.
PR 28-MAY-1999; 9905-0136392.
PR 01-JUN-1999; 9905-0136782.
PR 03-JUN-1999; 9905-0137222.
PR 04-JUN-1999; 9905-0137528.
PR 07-JUN-1999; 9905-0137502.
PR 08-JUN-1999; 9905-0137724.
PR 10-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 14-JUN-1999; 9905-0138847.
PR 16-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 17-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140053.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140823.
PR 28-JUN-1999; 9905-0140825.
PR 29-JUN-1999; 9905-0140891.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 02-JUL-1999; 9905-0142154.
PR 06-JUL-1999; 9905-0142055.
PR 08-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142820.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 19-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144125.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145324.
PR 26-JUL-1999; 9905-0145326.
PR 27-JUL-1999; 9905-0145313.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.


```

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157754.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159429.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159648.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161461.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 72.78; Score 2648; DR 21; Length 648.
Best Local Similarity 83.4%; Pred No 1 6e-203;
Matches 514; Conservative 43; Mismatches 57; Indels 2; Gaps 2.

QY 93 MESMNGCTISYCTTTCGATSHRRKQGAIOKEILRPLNACIPNCGT-ETSHLPKSGAT 151
DB 1 MESMNRGTDSYGTGTGATSHRRKQGAIOKEILRPLNACIPNCGT-ETSHLPKPTT 60

QY 152 RAAMIVRINILLDGYSSTFELEIATKPLNNNTPTCLPRTGTASRDLVLSIATLL 211
DB 61 RAAMIVRINILLDGYSSTFELEIATKPLNNNTPTCLPRTGTASRDLVLSIATLL 120

QY 212 TRPPNSKAVGPT-LEVLNAEKATAAV-VEGTFEELQFKRELVAN-NAVSSMAVLEFA 271
DB 121 TRPPNSKAVGPT-LEVLNAEKATAAV-VEGTFEELQFKRELVAN-NAVSSMAVLEFA 179

QY 272 NVLALISFVLSAIFARVMGCKPFTDHLTHKIKHHGCTLEAATATVYLDGSDYYKAOK 331
DB 180 NVLALISFVLSAIFARVMGCKPFTDHLTHKIKHHGCTLEAATATVYLDGSDYYKAOK 239

QY 332 VHEMDPLQKRPDRAVLRTPQWLRPQTEVIRPSTKMTFFPTNSVNDPLIVSNKALH 391
DB 240 VHEMDPLQKRPDRAVLRTPQWLRPQTEVIRPSTKMTFFPTNSVNDPLIVSNKALH 299

QY 392 GGNVCGTPTIVSMDNTRLAIATGKIMFAGQSELVNDEFNNKLEPSSLSGGRNPSLDYGRK 451
DB 300 GGNVCGTPTIVSMDNTRLAIATGKIMFAGQSELVNDEFNNKLEPSSLSGGRNPSLDYGRK 359

QY 452 GGIEMASYSCTFIPLANVININVOGAPQHNQDVNSIGLISAKTIAFAVDLLKIMSSYIL 511

```

```

DB 360 GAIEMASYSCTFIPLANVININVOGAPQHNQDVNSIGLISAKTIAFAVDLLKIMSSYIL 419
QY 512 VALGYSIDLKHLEENKSTVKNVSOVAKVLLIKVNVNHELHPSKFEKEDLLRVDPREVF 571
DB 420 VALGYSIDLKHLEENKSTVKNVSOVAKVLLIKVNVNHELHPSKFEKEDLLRVDPREVF 479
QY 572 AYIDVCGCTPTIVSMDNTRLAIATGKIMFAGQSELVNDEFNNKLEPSSLSGGRNPSLDYGRK 641
DB 480 AYIDVCGCTPTIVSMDNTRLAIATGKIMFAGQSELVNDEFNNKLEPSSLSGGRNPSLDYGRK 539
QY 632 VRIAYNDPLISTIPNRIKACRSYPLRYFVEELQCHITGQKVTSPGEPDHPVTAMKGO 691
DB 540 VRIAYNDPLISTIPNRIKACRSYPLRYFVEELQCHITGQKVTSPGEPDHPVTAMKGO 599
QY 692 IIDPLEELGGMNGEP 707
DB 600 IIDPLEELGGMNGEP 615

RESULT 13
AAG21947
ID AAG21947 standard; Protein: 423 AA.
AC
AC AAG21947;
XX
DI 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24692.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
CS Arabidopsis thaliana.
PN
PN BP1033405-A2.
PD
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301449
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 04-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

```

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142940.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148654.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150565.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160914.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160984.
PR 25-OCT-1999; 99US-0161484.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161339.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999: 9905-0161992.
PR 28-OCT-1999: 9905-0161993.
PR 29-OCT-1999: 9905-0162142.

Query Match 51.98; Score 1892; DB 21; Length 423,
Best Local Similarity 83.78; Pred. No. 3, 9e-143;
Matches 354; Conservative 33; Mismatches 36; Indels 0; Gaps 0;

OY 289 MOSKUPFHDTLHKLKHNGQIAAAIMCYLLDGSNYKAAQKHYHMYLOKRYODRAL 348
DB 1 MOSKUPFHDTLHKLKHNGQIAAAIMCYLLDGSNYKAAQKHYHMYLOKRYODRAL 60
OY 349 KTSQWMLSFQIEYRSTKMEKEINSVNDPLIVSRKALHSGNFQSTPFGVSMONT 408
DB 61 KTSQWMLSGQIEYRSTKMEKEINSVNDPLIVSRKALHSGNFQSTPFGVSMONT 120
OY 409 LAIAIAIGKIMFQFSLVNDYFNNGIIPSMISGRRNPISLDYCKGKHJAMASYCELOPLA 468
DB 121 LAIAIAIGKIMFQFSLVNDYFNNGIIPSMISGRRNPISLDYCKGKHJAMASYCELOPLA 180
OY 469 NPVTNVQSAFOHNDPVNSIGTISAKTIAEAVDILKMSSTYLVAMIGOSIDIRHLEHMK 528
DB 181 NPVTNVQSAFOHNDPVNSIGTISAKTIAEAVDILKMSSTYLVAMIGOSIDIRHLEHMK 240
OY 529 STYKNTVSQVAKKVLTMVNSDELHPSRPTEKDLRVPVPEVFAVTECVNSTIPLMKL 588
DB 241 STYKNTVSQVAKKVLTMVNSDELHPSRPTEKDLRVPVPEVFAVTECVNSTIPLMKL 300
OY 589 KOVLVNDHALNCHTEKNTNTSIFOKIATFHEHLKVLLEKVGAVIAEHNDLSIPNPK 648
DB 301 KOVLVNDHALNCHTEKNTNTSIFOKIATFHEHLKVLLEKVGAVIAEHNDLSIPNPK 360
OY 649 ACSQVPLVYFVREHORGFTGCKKVTSPGEFEDRYFTAMCKOUIIDPLECIGWNCPL 708
DB 361 ACSQVPLVYFVREHORGFTGCKKVTSPGEFEDRYFTAMCKOUIIDPLECIGWNCPL 420
OY 709 PTC 711
DB 421 PTC 423

RESULT 14
AAB16402
ID AAB16402 standard; Protein: 424 AA.
XX
AC AAB16402;
XX
DT 24-OCT-2000 (first entry)
XX
DE Eucalyptus grandis PAL protein sequence SEQ ID NO. 423
XX
KW plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;
KM pinus radiata; Monterey pine.
XX
OS Eucalyptus grandis.
XX
PN WO200022099-A1.
XX
PD 20-APR-2000.
XX
PF 06-OCT-1999; 9905-0200168.
XX
PK 09-OCT-1998; 9905-0164789
XX
PK 14-JUL-1999; 9905-0143811.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Hloksberg LM, Hanykkaia LJ;
XX
WM 2000-317962/27.
XX
PT Novel polynucleotide encoding enzymes involved in lignin-biosynthetic

PT pathway useful for producing transgenic plants especially eucalyptus
PT and pine species having altered lignin content, composition and
PT structure

PS Claim 18, Page 170-171, 213pp, English.

XX The present invention describes isolated polynucleotides and proteins
XX encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),
XX cinnamate 3-hydroxylase (C3H), phenolase (PHL), o-methyl transferase
XX (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase
XX (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate-CoA ligase (4CL),
XX coniferin glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),
XX laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,
XX caffeic acid methyl transferase, caffeoyl CoA methyl transferase,
XX cinnamoyl transferase, cytochrome P450 1XX1A, diphenol oxidase, flavanol
XX glucanase, flavanoid hydroxylase, and isoflavone reductase,
XX which are involved in the lignin biosynthetic pathway. The
XX polynucleotides can be used for modulating lignin content, lignin
XX composition and the structure of a plant, especially eucalyptus and pine
XX species, and for modifying the activity of an enzyme involved in lignin
XX biosynthetic pathway, and for producing a plant having altered lignin
XX content, composition and structure. They can be used for designing probes
XX and primers useful for detecting similar DNA and RNA sequences in any
XX organism and for PCR amplification. The lignin content can be efficiently
XX modified using the polynucleotides. AAB7998 to AAB6201 and AAB1641 to
XX AAB16449 represent the polynucleotide and protein sequences used in the
XX exemplification of the present invention.

SO Sequence 424 AA:

Query Match 49.4%; Score 1799; DB 21; Length 424;
Best Local Similarity 82.68; Pred. No. 1, 1e-135;

Matches 351; Conservative 37; Mismatches 29; Indels 8; Gaps 2;

OY 1 MEMGNVNGVNNELCT-----KDIPLMNGVAAALATGSHLDYKKKVAAPKRPVYKIGCT 55
DB 3 MESTTGTGMLSLSLVNAAGSHADPLNMAAAGATGSHLDYKRPVEEYRPAVPLVLES 62
OY 56 LTVSVAAGIAAANDSVFVVELSEAKACGKASSIWMMSMKGIDUSGVYDGCATSHR 115
DB 63 LITLAQVAAYAS---QHGCVELHSPAARKRPVKASSIWMMSMKKIDUSGVYDGCATSHR 119
OY 116 PTKGSAALAKELIPUNASIPNNTEITSHLEPSAIPAAIVPINTLLGYSGITFEILE 175
DB 120 RTKGCGALQKELIRPUNAGIPNGTSCHTLPQSSIPRAAMIVRVNILLGYSGITFEILE 179
OY 176 AITKPLNNNTTQPLRGCTTASGDIYPISTYAGLITGSPNSKAAGPTGCVINAKKAA 235
DB 180 AITKPLNNNTTQPLRGCTTASGDIYPISTYAGLITGSPNSKAAGPTGCVINAKKAA 239
OY 236 ACVCGFPELOHKEGIALVNGTAVSGMASVLPANVIALISEVLSAIFAEVWQKPEF 295
DB 240 AGIDTGFPELOHKEGIALVNGTAVSGIASIVLPANVIALISEVLSAIFAEVWQKPEF 299
OY 296 TDHLTHKLKHHPEQTEAAIMEYILLGSDYVKAQKVIHEMPELOKPKODRYALRTSPQWL 355
DB 300 TDHLTHKLKHHPEQTEAAIMEYILLGSDYVKAQKRIHEMPELOKPKODRYALRTSPQWL 359
OY 356 GICQLEVRSSIKMIEKINSVNDPLIVSRKALHSGNFQSTPFGVSMONTIRLAAAG 415
DB 360 GICQLEVRSSIKMIEKINSVNDPLIVSRKALHSGNFQSTPFGVSMONTIRLAAAG 419
OY 415 KLMFA 420
DB 420 KLMFA 424

RESULT 15
AAB78982
ID AAB78982 standard; Protein: 495 AA.
XX
AC AAB78982;
XX
PT

GENERAL INFORMATION:
APPLICANT: Yoshida, Roberta
TITLE OF INVENTION: Phenylalanine Ammonia lyase Polypeptide and
TITLE OF INVENTION: Polypeptide Sequences and Methods of Obtaining and
TITLE OF INVENTION: Using Same
FILE REFERENCE: 29479/500NSC
CURRENT APPLICATION NUMBER: US/09/624.693A
CURRENT FILING DATE: 2000-07-24
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 19
LENGTH: 716
TYPE: PRT
ORGANISM: Rhodotorula toruloides
US-09-624-693A-19

Query Match 27.3%, Score 994.5, Dh 4, Length 716,
Best Local Similarity 35.9%, Pred. No. 4,9e-81,
Matches 270; Conservative 123, Mismatches 250, Indels 109, Gaps 23;

QY 8 NGVNNELCTKIDPLNMCVAAEALTGSHL-----DEKKVVAERFRKVVYKLGSETLTVS 59
DB 14 NGVAS---AKQAVNGASINLAAGSHLPTQVTVQVDIVKMLAAPDSTELIDGYSLINIG 70
QY 60 QVAGIAANDSDTVKVELSEARAGYKASSDMVMSNMKCTDSCVGTTCGATSHRRTKO 119
DB 71 DV--VSAARKGRPVYRKSDDELRSKIDKSVETLSQLSMSV--YVTTGPGASATLPTED 126
QY 120 GCAIQELIRPLINAGI-----FCNGTETSHLIPHSATRAAMIVINILIOGSGIR 170
DB 127 AISTQALLENHOLGCVLPSFISFRLGRLENS--LPLEVVGKAMTIVNLSILKCHSAVR 184
QY 171 FELLEATFELNNITPCLPLRGITIASGDLVPLSYAGLLTGPNNSKA-VGPG--EVL 227
DB 185 LVVLEATFELNNITPCLPLRGITIASGDLVPLSYAGLLTGPNNSKA-VGPG--EVL 227
QY 228 NAKKAFVAGVGEFPELOPKKGLALVNGTAVSGMASMVLFDANVLAISLSEVLSALFAE 287
DB 245 YAREMALNLEF--VLGPKKGLVNGTAVSASMATLALHDAMLSLSOSTTAMTIVE 302
QY 288 VMOGRP-EFTDHLTHKLIKHPGQIEAALIMFYILDGSDY-VKAAQKVFHMPDQPKODR 345
DB 303 AMVGHGSPFHPFLHCVTPRHPTQIEVACNIRKLLEGSFPAVHHEFFVVKDEESILRQDR 362
QY 346 YALPTSPOMLGPDI-EVIPSSTKMIERELNSVNDPLIDVSNKALHGCNFGTPTIVSM 404
DB 363 YPLPTSPOMLGPDI-VLIPSSSTKMIERELNSVNDPLIDVSNKALHGCNFGTPTIVSM 422
QY 405 DNTFLAIAAIGKIMFAUSELVNDFYNNGLPSNLSGCKNPSIDYGFKGGELIAMAISYCSSEL 464
DB 423 EKTRGLAIGKINFTQITLMLNAGNNRGLPSCLA-AEDPISLHYCKGLIDIAAAAYTSEL 481
QY 465 QPLANVTVNHVOSAHQNDVNSIGLISAKRTAFVADIIKIMSSYVLAICOSTIDLRHLE 524
DB 482 GHLANVTVNHVOPAEVANAQVNSIALISAKRTTSSNDVLSILATHLICVLAIDILRAIE 541
QY 525 ENNKSTVKNTVSQVAKKVLITMGVNGELHPSRCEKDLRVVDREVFAVTDVCSGTPL 584
DB 542 FPKKK-----QFGPA-----IVSLIDQHP-----GSAMT 565
QY 585 MOKLROVLVDHALNNGETKNTNTSIFOKIATFEELKVLIPKEVEGY----- 632
DB 566 GSNLKRDELVE-KVNTKLARLEDTNSYDLPVPMIDAFSEMASTVEVLSSTLSLAAVNA 624
QY 633 -RIAYENDITLILNRIK-----ACRSYP-----LYRFRRELIG--RG--FI 668
DB 625 WKXAAAFSAISLIRKVRHETFWASASTSYAISLIRKTVILAFVKRIHJOVKARKQGYFI 684
QY 669 TGEKVTSPDEEDRVFTAMCKGQIIDPLLECL 700
DB 685 GKOFVTL-GSNVSKLYEALIKGSHINNVLKML 715

RESULT 5
US-09-627-216A-8
Sequence 8, Application US/09627216A
Patent No. 6368837

GENERAL INFORMATION:
APPLICANT: Sartorius, Sime F
APPLICANT: Tang, Xiao-Song
APPLICANT: Qi, Wei-Wei
APPLICANT: Vannelli, Todd
APPLICANT: Galenby, Anthony
TITLE OF INVENTION: Bioproduction of para-Hydroxycinnamic Acid
FILE REFERENCE: H01009 US NA
CURRENT APPLICATION NUMBER: US/09/627.216A
CURRENT FILING DATE: 2000-07-27
PRIOR APPLICATION NUMBER: 60/147,719
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Microsoft Office 97
SEQ ID NO: 8
LENGTH: 716
TYPE: PRT
ORGANISM: Rhodotorula glutinis
US-09-627-216A-8

Query Match 27.3%, Score 994.5, Dh 4, Length 716,
Best Local Similarity 35.9%, Pred. No. 4,9e-81,
Matches 270; Conservative 123, Mismatches 250, Indels 109, Gaps 23;

QY 8 NGVNNELCTKIDPLNMCVAAEALTGSHL-----DEKKVVAERFRKVVYKLGSETLTVS 59
DB 14 NGVAS---AKQAVNGASINLAAGSHLPTQVTVQVDIVKMLAAPDSTELIDGYSLINIG 70
QY 60 QVAGIAANDSDTVKVELSEARAGYKASSDMVMSNMKCTDSCVGTTCGATSHRRTKO 119
DB 71 DV--VSAARKGRPVYRKSDDELRSKIDKSVETLSQLSMSV--YVTTGPGASATLPTED 126
QY 120 GCAIQELIRPLINAGI-----FCNGTETSHLIPHSATRAAMIVINILIOGSGIR 170
DB 127 AISTQALLENHOLGCVLPSFISFRLGRLENS--LPLEVVGKAMTIVNLSILKCHSAVR 184
QY 171 FELLEATFELNNITPCLPLRGITIASGDLVPLSYAGLLTGPNNSKA-VGPG--EVL 227
DB 185 LVVLEATFELNNITPCLPLRGITIASGDLVPLSYAGLLTGPNNSKA-VGPG--EVL 227
QY 228 NAKKAFVAGVGEFPELOPKKGLALVNGTAVSGMASMVLFDANVLAISLSEVLSALFAE 287
DB 245 YAREMALNLEF--VLGPKKGLVNGTAVSASMATLALHDAMLSLSOSTTAMTIVE 302
QY 288 VMOGRP-EFTDHLTHKLIKHPGQIEAALIMFYILDGSDY-VKAAQKVFHMPDQPKODR 345
DB 303 AMVGHGSPFHPFLHCVTPRHPTQIEVACNIRKLLEGSFPAVHHEFFVVKDEESILRQDR 362
QY 346 YALPTSPOMLGPDI-EVIPSSTKMIERELNSVNDPLIDVSNKALHGCNFGTPTIVSM 404
DB 363 YPLPTSPOMLGPDI-VLIPSSSTKMIERELNSVNDPLIDVSNKALHGCNFGTPTIVSM 422
QY 405 DNTFLAIAAIGKIMFAUSELVNDFYNNGLPSNLSGCKNPSIDYGFKGGELIAMAISYCSSEL 464
DB 423 EKTRGLAIGKINFTQITLMLNAGNNRGLPSCLA-AEDPISLHYCKGLIDIAAAAYTSEL 481
QY 465 QPLANVTVNHVOSAHQNDVNSIGLISAKRTAFVADIIKIMSSYVLAICOSTIDLRHLE 524
DB 482 GHLANVTVNHVOPAEVANAQVNSIALISAKRTTSSNDVLSILATHLICVLAIDILRAIE 541
QY 525 ENNKSTVKNTVSQVAKKVLITMGVNGELHPSRCEKDLRVVDREVFAVTDVCSGTPL 584
DB 542 FPKKK-----QFGPA-----IVSLIDQHP-----GSAMT 565
QY 585 MOKLROVLVDHALNNGETKNTNTSIFOKIATFEELKVLIPKEVEGY----- 632
DB 566 GSNLKRDELVE-KVNTKLARLEDTNSYDLPVPMIDAFSEMASTVEVLSSTLSLAAVNA 624

QY 633 -RIAYENDTLSPNRK-----ACRSYP-----LYRPREELG---RG--FL 668
 Db 625 WKVAAAEASISLTRQVRETFEWSAASTSSPALSTYSPRTQILLYAVREBELGAKARQGVFL 684
 QY 669 TGEKVTSPGEEDRVFTAMCKGQIIDLPLECL 700
 Db 685 GKQEVTL-GSNVSKIYEAIKSGRINVLKML 715

RESULT 6

US-09-627-216A-10
 ; Sequence 10, Application US/09627216A
 ; Patent No. 6368837
 ; GENERAL INFORMATION:
 ; APPLICANT: Sartisiani, Sima F
 ; APPLICANT: Tang, Xiao-Song
 ; APPLICANT: Qi, Wei Wei
 ; APPLICANT: Vannelli, Todd
 ; APPLICANT: Gatenby, Anthony
 ; TITLE OF INVENTION: Bioproduction of para-Hydroxycinnamic Acid
 ; FILE REFERENCE: BC1009 US NA
 ; CURRENT APPLICATION NUMBER: US/09/627,216A
 ; CURRENT FILING DATE: 2000-07-27
 ; PRIOR APPLICATION NUMBER: 60/147,719
 ; PRIOR FILING DATE: 1999-08-06
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 10
 ; LENGTH: 716
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: mutant from
 ; OTHER INFORMATION: Rhodotorula glutinis
 US-09-627-216A-10

Query Match 27.2%; Score 991.5; DB 4; Length 716;
 Best Local Similarity 35.9%; Pred. No. 9.2e-81;
 Matches 270; Conservative 122; Mismatches 251; Indels 109; Gaps 23;

QY 8 NGVVELCICKDPLMNGVAFAALTCGSHL-----DEYKKVAFRRKPVYKIGETITVS 59
 Db 14 NGVAS---AKQAVNGASINLAVASHLPTTYQVDIYEKKLAAPTSITELGYSINTLG 70
 QY 60 QVAGIAAANDSTVYKVELSEARAGVAKASPDWVMSMKGTDSTGYTGTGATSHRRTKQ 119
 Db 71 DV--VSAARKGRPVYKDSDELRSKIDKSYEFLRSQLSMSV--YGVTTGCGSADTRTED 126
 QY 120 GGALCKELIRFLNAGI-----FGNGTETSHTPHSATRAMIYRINTLLQGYSGIR 170
 Db 127 AISLCKALLEHQLGCVLPSSPDSFRLRGLEN--LPLEVYKGMATIRVNSLTRGHSAYR 184
 QY 171 FEILEATKFLNNNTIPCLPLRGITITASGLVPLSYIAGLLTGRPNKA--VGPTG--EVL 227
 Db 185 LVVLEALNLEINLHGITRIPVLRGITSASGDLSPSTYIAAISGHPDKYVHVHEGKKIL 244
 QY 228 NAEKAPAAAGVGGFFELQPKKGLALVNGTAVSGMASMVLFDANVALLSEVLSAIFAE 287
 Db 245 YAREMAALFENLR--VVLGPKKGLVNGTAVSASMATLALHDAMHLSLISOSLTAATVE 302
 QY 288 VVGQKP--ETDHLTHLKHHPGQIEAAIMEYILDGSDY--VKAQKQHEMDLPQKQDR 345
 Db 303 AMVGAGSGFHPFLHDVTRPHPTQIEVACNIRKLLGSRFAVHHEEVKVKVADDEGILLKQDR 362
 QY 346 VALKTSQWLPQPI--EVIIRSTKMIEREINSVNDPLIDVSRNKALGCGNFOGTPIGVSM 404
 Db 363 YLRTSPQWLGFLVSDLIHAHVLITIEAGSTTDNPLIDENKSHSGNFGVAAVANTM 422
 QY 405 DNTRIATAIGKLMFAQSELYNDVFNNGCLPSNLISGRNPSLDYGFKGGELIAMAATSCSEL 464
 Db 423 ETRTGLAQICKLNTQJLEMLNAGMNGDPSCLA--AEDPSLSYHCKGIDIAAAYATSEL 481
 QY 465 QELAPVTNIHVSAEJHNUQVNSISGLISARKTAEAVDILKLMSTYIVLVCQSIDLRHIE 524

Db 482 GHLANPVTHQAPAMAGVANSALISARTTSSNOVLSILLATHLYCVLQADIDRATE 541
 QY 525 ENMKSTVKNVTSQVAKKVLTWGVNGEILHPSRFCKDLLRVVDREYRAYIDVCGSTYPI 584
 Db 542 FEFKK-----QEGPA-----IVSLIDQHF-----GSAMT 565
 QY 585 MQKRLQVLDHALNNGETEKNTNISFQKIAFEELKVLIPKEVEGV-----632
 Db 566 GSNLDELVE--KVNKTIAKRLQJNSTDVLPBWHDAFSPAAGVVEVLSLAAVNA 624
 QY 633 -RIAYENDTLSPNRK-----ACRSYP-----LYRPREELG---RG--FL 668
 Db 625 WKVAAAEASISLTRQVRETFEWSAASTSSPALSTYSPRTQILLYAVREBELGAKARQGVFL 684
 QY 669 TGEKVTSPGEEDRVFTAMCKGQIIDLPLECL 700
 Db 685 GKQEVTL-GSNVSKIYEAIKSGRINVLKML 715

RESULT 7

US-09-624-693A-17
 ; Sequence 17, Application US/09624693A
 ; Patent No. 6353468
 ; GENERAL INFORMATION:
 ; APPLICANT: Yoshida, Roberta
 ; APPLICANT: Kootstra, Anna
 ; TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
 ; TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
 ; FILE REFERENCE: 29479/500NSC
 ; CURRENT APPLICATION NUMBER: US/09/624,693A
 ; CURRENT FILING DATE: 2000-07-24
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 17
 ; LENGTH: 713
 ; TYPE: PRT
 ; ORGANISM: Rhodotorula mucilaginosa
 US-09-624-693A-17

Query Match 26.7%; Score 971.5; DB 4; Length 713;
 Best Local Similarity 36.7%; Pred. No. 5.9e-79;
 Matches 267; Conservative 125; Mismatches 277; Indels 59; Gaps 22;

QY 1 MENGHNHGVNVEICL-KDPLMNGVAFAALTCGSHLDEYKKVAFRRKPVYKIGETITVS 58
 Db 16 LSNGLHAAANAGDGVHKKKTAGAGSLPTTETTYQVDIYERILADAGTQIKIDGTYTLL 75
 QY 59 SOVAGIAAANDSTVYKVELSEARAGVAKASPDWVMSMKGTDSTGYTGTGATSHRRTK 118
 Db 76 GDVVG--AARGRGVKAVKADSPHIREKIDASYEFLRQLDMSV--YGVTTGCGSADTRTE 131
 QY 119 QGGALEKELIRFLNAGI-----FGNGTETSHTPHSATRAMIYRINTLLQGYSGI 169
 Db 132 DAISLCKALLEHQLGCVLPSSPDSFRLRGLEN--LPLEVYKGMATIRVNSLTRGHSAY 189
 QY 170 FEILEATKFLNNNTIPCLPLRGITITASGLVPLSYIAGLLTGRPNKA--VGPTG--EVL 229
 Db 190 RIVVLEALNLEINLHGITRIPVLRGITSASGDLSPSTYIAAISGHPDKYVHVHEGKKIL 248
 QY 230 EKAFAAGVGGFFELQPKKGLALVNGTAVSGMASMVLFDANVALLSEVLSAIFAEY 289
 Db 249 QEALAKGLQPI--VVLGPKKGLVNGTAVSASMATLALHDAMHLSLISOSLTAATVE 306
 QY 290 QGKP--ETDHLTHLKHHPGQIEAAIMEYILDGSDY--VKAQKQHEMDLPQKQDR 347
 Db 307 VGHAGSFHPFLHDVTRPHPTQIEVACNIRKLLGSRFAVHHEEVKVKVADDEGILLKQDR 366
 QY 348 LRTSPQWLPQPI--EVIIRSTKMIEREINSVNDPLIDVSRNKALGCGNFOGTPIGVSM 406
 Db 367 LRSPQWLGFLVSDLIHAHVLITIEAGSTTDNPLIDLEKKMHHGATMASSVGTMEK 426

```

QY 407 TRLAIAIGKLMFAVSELVNFYNNLPSLNSGIRNPSTIYVFKGJEIAMASGSELD 466
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 427 TRIAVAKIKKRVSTQVIFHMLNAGMNNALPSCIA-AEDISLSTHCKGLDIAAAATSEIGH 485
QY 467 LANPTNHVOSAKOHNOVNSIGLISAKRTARAVDILKIMSTYVALCOSIDIRHLREN 526
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 486 LANPSTHVOAPAMNSQVAINSLALISARTAEANDVLNLLATILHYVLQVAVDEBRAMEFE 545
QY 527 MKSTVNTVSVQAKKYLIMGVNGBLHPSHCKKLLKVVDRKYVATIDVCSGTIYLMQ 586
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 546 HKAPFPMVTEILKQHF-----GALATAHVPK-----VRKSTYKRIQO--NNSYDLEQ 592
QY 587 KLRQVL-----VDIALNGSETEKNTNTSIFOKIATPEEELKVLBPKEVEVRIAYENDT 640
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 593 RHMDTSVATGAVVAIAAQEVSILASIA--WKVACAKKALAL-----TRSVKDSYMAP 645
QY 641 LSLPNRIK--ACHSYLYKPVHREIG---KG--FLTCRKVTSNCEPDEHVFAMCKQOI 692
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 646 SSSSPALKYLSPTFVLVSFVREEVVKAHFGCVLTIKGEVIT GINVSIFETAKISRL 704
QY 693 IDPLILPCL 700
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 705 APVILVKKM 712

```

RESULT 8

```

US-09-624-693A-21
: Sequence 21, Application US/09624693A
: Patent No. 635468
: GENERAL INFORMATION:
: APPLICANT: Yoshida, Roberta
: APPLICANT: Koolstra, Anna
: TITLE OF INVENTION: Phenylalanine Ammonia lyase polypeptide and
: TITLE OF INVENTION: Polynucleotide sequences and methods of obtaining and
: TITLE OF INVENTION: Using Same
: FILE REFERENCE: 29479/500NSC
: CURRENT APPLICATION NUMBER: US/09/624-693A
: CURRENT FILING DATE: 2000-07-24
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 21
: LENGTH: 726
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: SITE
: LOCATION: 112, 16, 17, 20, 24, 25, 28, 30, 38, 42, 47, 48, 56, 57, 62, 66,
: OTHER INFORMATION: "Xaa" means any amino acid; "Xaa" means no consensus at that
: OTHER INFORMATION: position
: OTHER INFORMATION: Description of Artificial Sequence: Consensus of
: OTHER INFORMATION: SEQ ID NOS: 13, 17, and 19
US-09-624-693A-21

```

Query Match

26.5%: Score 967; DP 4; Length 726;

Best Local Similarity

36.9%: Pred. No. 15e-78;

Matches 263; Conservative 108; Mismatches 262; Indels 80; Gaps 20;

```

QY 32 SHIDFVKKKMAF-FKKPVKICQETITVSOVAGIAAANDSDVKVSEAAAGYKASSD 90
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 49 TOLDIYFKXKAPXTDDXXELGYSLTLDVYG--AAKRGVRRVXDXDFTKXLDKQSE 106
QY 91 WVMESNKKGTDSYVTTGRTATSHRFTKQJALQKELIFELNAGI-----FGNGTE 141
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 107 FLKXQJXNSV--YGVITFGQASADIRTEIDALISLOKALIEHQICVLPISXDSFXGLRLK 164
QY 142 TSTIDLSHATRAMVIRINTLLOGYSIRFELLEITKTLNNTTPCPPLRFTTASGDL 201
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 165 NS--LPLEVYFGAMTIRVNSITRHSANVLVLEALTNLNGITPIVPLRGTISASGDL 222
QY 202 VPLSYIAGLITGPNNSKA--VGPFG--FVILNAKKAPAAACVPGCPPELQPKKGLAVNGTA 258
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 223 SPLSYIAAATITFPGSKVIVHIEGXKIKAKHFAIALFLER--VLLPKKELILVNITA 280

```

```

QY 254 VSGMASMVLPTANVALLSEVLSAIAFVWGKRP-ETTHLTIKRLKHQGTDAIME 317
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 281 VSASMTATLAIHAAHMLISLSOAITALTIVAMYGAGSHHPLIHVTRDPQIIVANRIK 340
QY 318 YLLDSUDY--VKAQKKVHEMDPLQKKQDURAYLRTSPQWICQI--EVIKSTKMLEPINS 375
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 341 TLLESXKFAVHIEEEKVAKDDEGILKQDQYPLKTSFQWLSPLVNSMILIAVLELENQS 400
QY 376 VNDRPLIVDSRNKALHCGHPCSTPIGVSMONTKILAAIAIKIMAGSELVNDVYNNQIP 435
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 401 TIDNPILIDENKXTHHGGNFGASAVXNTEKTRIALALIGKILNFQIEMENAGMNNKQIP 460
QY 436 SNLSGGRNSLIDYGFGRITAMASGSELDPLANVTNHWGSAELHNDVNSLISARK 495
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 461 SCIA-AEDISLSTHCKGLDIAAAATSEIGHLANVTTIHVOIAHNGOAVNSIALISAR 519
QY 496 TAAVADILKIMSTYVALCOSIDIRHLRENKSTVKNVSOAKKYLIMGVNGBLHPS 555
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 520 IAAVAVLSLLATILHYVLQVAVDEBRAMEFEKKQFQXXXXLXQIF---GXKXTX 574
QY 556 FCKKDLIRVYDRKYVATIDVCSGTIPLMOKLROYLVDAIANGETKNTNTSIFOKIA 615
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 575 XXXXELXXKVVXX-----LKKRLQOT-----NSYDLRPMHDAFSXATG 613
QY 616 IFEEELK-----VLLPEVEGVRIAYENDLSIPRIK-----ACKSY 653
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 614 TVEXLSSXXXXXVSLAAVANAKVAAKKAISLTXXVXXXXXXAISSSPALXLSPTIK 673
QY 654 PLVRFVREIG---KG--FLTCRKVTSNCEPDEHVFAMCKQOIIDPLILPCL 700
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 674 VLVSFVREELGVAPRGVFLKQEVIT--GSNVSIFETAKISGPIXLVYML 725

```

RESULT 9

```

US-09-615-192A-327
: Sequence 327, Application US/09615192A
: Patent No. 6410718
: GENERAL INFORMATION:
: APPLICANT: Bloksberg, Leonard N.
: APPLICANT: Havukkala, Jukka
: TITLE OF INVENTION: Materials and Methods for the
: TITLE OF INVENTION: Modification of Plant Lignin Content
: FILE REFERENCE: 11000 1003c40
: CURRENT APPLICATION NUMBER: US/09/615-192A
: CURRENT FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 08/975,316
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: US 08/714,000
: PRIOR FILING DATE: 1996-04-11
: PRIOR APPLICATION NUMBER: US 08/164,789
: PRIOR FILING DATE: 1998-10-09
: NUMBER OF SEQ ID NOS: 405
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 327
: LENGTH: 226
: TYPE: PRT
: ORGANISM: Eucalyptus grandis
US-09-615-192A-327

```

Query Match

25.8%: Score 939; DP 4; Length 226;

Best Local Similarity

80.5%: Pred. No. 6.9e-77;

Matches 182; Conservative 19; Mismatches 25; Indels 0; Gaps 0;

```

QY 444 PSLDYGKQKGLIAMASGSELDPLANPVNHNHVSAMHNOVNSIGLISAKRTARAVDIL 503
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 1 PSLDYGKQKGLIAMASGSELDPLANPVNHNHVSAMHNOVNSIGLISAKRTARAVDIL 60
QY 504 KMSSTYVALCOSIDIRLLENMSTVKNVSVQAKKYLIMGVNGBLHPSRFEKDLK 563
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 61 KMSSTYVALCOSIDIRLLENMSTVKNVSVQAKKYLIMGVNGBLHPSRFEKDLK 120
QY 564 VLVREVFVYLLGVNSGITPIPLMQLROYLVDAIANGETKNTNTSIFOKIATPEEELK 623
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

```

DB 121 VVREVEFAYIDUP:SAIYPLMOKLROVLVDALLDVEKNEPSTISIFUKIGAEFEELKA 180
 QY 624 ILPKVEGVRIAYENDULSTIPNRIKACRSYPIYFVVEHIGRGUJT 669
 DB 181 LIPKVEFNARACPFESGNSAIANKIRGCRSYPLIFVVEHIGRGUJT 226

RESULT 10
 US-09-624-693A-15
 : Sequence 15, Application US/09624693A
 : Patent No. 6355468
 : GENERAL INFORMATION:
 : APPLICANT: Yoshida, Roberla
 : APPLICANT: Kootstra, Anna
 : TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
 : TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
 : TITLE OF INVENTION: Using Same
 : FILE REFERENCE: 29479/500NSC
 : CURRENT APPLICATION NUMBER: US/09/624.693A
 : CURRENT FILING DATE: 2000-07-24
 : NUMBER OF SEQ ID NOS: 25
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 15
 : LENGTH: 740
 : TYPE: PRT
 : ORGANISM: Amanita muscaria
 US-09-624-693A-15

Query Match 23.2%, Score 844.5; DB 4; Length 740;
 Best Local Similarity 31.3%; Pred. No. 2e-67;
 Matches 229; Conservative 153; Mismatches 266; Indels 83; Gaps 21;

QY 19 PLNGVVAADALT-----GSHDEVKKVAAPFRKPVYKLGSEITLVSOVAGIAA 66
 DB 27 PVN-GFKATALSKASRTMTKTSALSOFLAYELECYKNGRAIKVDQDTLSIAVA--AA 83
 QY 67 ANDSDTVVELSEARAGVYKASDWMVMKNGTDSYGVITGREGATSHRETKUGSALQKE 126
 DB 84 ARYNAVEIDESPVLKVEYKKSQALAIANKSTGASVGLSTGFGSGADTRDKRMILGFA 143
 QY 127 LIRPIMAGIFGNGTP-----TSHLLPHSATRAAMIVRINLLDQYSGIRFELIEA 176
 DB 144 LLDHGHVGLPSTPELVDLPDADNNTSMPEAMIRGAILIRMSLLIRHSGIRWELIEK 203
 QY 177 ITRKPLANNITPCLRGITITASGULVPISTYAGLLTGPNSKAV-GPT---CHVINAER 231
 DB 204 MEHLIAAVYIVVPIRGISISSGDLSPSYIAGTIIIGNPSIKYVHGPKSGQIRQIOSSKD 263
 QY 233 AFNAAGVEGFELOPKFEGIALVNGIANGSGMASVLFEDANVLLALSEVLSATFAEYMOG 291
 DB 264 VIALHNIEP--FPLESKREPLDILNGTASASVAAALALNEALHVLVLAQVCTAMQTHALIG 321
 QY 292 KPEFTDHLTHKLKHPQSIHAAIMEYILDGSDYKRAQ--KYHEM--DPLKPKADRY 346
 DB 322 TRASHAPFIHATARPHPGOVECAENINWILDGS---KLAEHEHEVLEDDKYLTRQDRY 378
 QY 347 AIRTSPOMLGQIIVIRSTKMIKEKIN--SVNPNLLIDVSKKMLHGCNQGIPICVSM 404
 DB 379 PLRTSPQELGQIDILISAFQTVQOECVYLPATDNPILIDGTGSHHGNGQAMAVYINAM 438
 QY 405 ENTKLALAIAGIKLMEAFSELDVFNNGSLPNSLGRKNSLDYGEFGKELIAMSASYCEL 464
 DB 439 EKTRIALAHVCKLILFSQSTELVNPAMNKGIPPSVA--ATPDSLNHAAGLIDIAAAYAE- 496
 QY 465 OFLAPVPIVNHQASADQHNQDYNISLISAKKTAFAVDILKLMSTVYVALCOSIDILRHE 524
 DB 497 --ATPGTHISQASMEHQAVNSLALISARATITSLSEVLTSILASYLICQALDLFALQ 553
 QY 525 ENMKSTVANTYSQVAKKVLIMGVNGELAPS--RPECMDLLRVVDREVFAYIDVYSGT- 581
 DB 554 REFLPG-----LDLIIIEELRSSGSLSEQOMKLOONLTSAFEDHLKITT 601
 QY 582 -----YPLMOKLROVLVDHALNNGETEKNTSIFOKIATFEELKVLPREVEGVRI 634

DB 602 MONTORMTMATSSSVLLQFETDGSASVPSSODLLSSVSSQSSVATRSSVLMODLRK 661
 QY 635 AT-ENHITISINWIKACRSYPIYFVVEHIG-----FGLTGT---KRVISPGHEED 681
 DB 662 EYTFSDPPTPASQYIRKTPPVYQFIPITIGVKKHSGENYKPFYNSLVEDVTI--GQNSI 720
 QY 682 RYETAMCKGOI 692
 DB 721 RYVESIRDOGM 731

RESULT 11
 US-09-615-192A-326
 : Sequence 326, Application US/09615192A
 : Patent No. 6410718
 : GENERAL INFORMATION:
 : APPLICANT: Bloksberg, Leonard N.
 : APPLICANT: Havukkala, Ilkka
 : TITLE OF INVENTION: Materials and Methods for the
 : TITLE OF INVENTION: Modification of Plant Lignin Content
 : FILE REFERENCE: 11000.1003c4U
 : CURRENT APPLICATION NUMBER: US/09/615.192A
 : CURRENT FILING DATE: 2000-07-12
 : PRIOR APPLICATION NUMBER: US 08/975.316
 : PRIOR FILING DATE: 1997-11-21
 : PRIOR APPLICATION NUMBER: US 08/713.000
 : PRIOR FILING DATE: 1996-09-11
 : PRIOR APPLICATION NUMBER: US 09/169.789
 : PRIOR FILING DATE: 1998-10-09
 : NUMBER OF SEQ ID NOS: 405
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 325
 : LENGTH: 164
 : TYPE: PRT
 : ORGANISM: Eucalyptus grandis
 US-09-615-192A-326

Query Match 17.9%; Score 653; DB 4; Length 164;
 Best Local Similarity 73.6%; Pred. No. 3e-51;
 Matches 120; Conservative 16; Mismatches 27; Indels 0; Gaps 0;

QY 548 NCELHPSRCECKDILRYVDREYFAYIDVCSYIYPLMOKLROVLVDHALNNGETEKNTN 607
 DB 1 NCELHPSHYCERDILKVDYDFRHFVYANDGSAIYPLMOKLROVLVDQALVNGESLINS 60
 QY 608 TSIFQKIATFEELKVLILPKVEGVRIAYENDULSTIPNRIKACRSYPIYFVVEHIGRGUJT 667
 DB 61 TSIFQKIATFEELKVLILPKVEGVRIAYENDULSTIPNRIKACRSYPIYFVVEHIGRGUJT 667
 QY 668 LTGKVTSPGEEFDRVFTAMCKGOIIDLPLECLGANGGEPLPI 710
 DB 121 LTGKVTSPGEEFDRVFTAMCKGOIIDLPLECLGANGGEPLPI 163
 RESULT 12
 US-09-615-192A-326
 : Sequence 326, Application US/09615192A
 : Patent No. 6410718
 : GENERAL INFORMATION:
 : APPLICANT: Bloksberg, Leonard N.
 : APPLICANT: Havukkala, Ilkka
 : TITLE OF INVENTION: Materials and Methods for the
 : TITLE OF INVENTION: Modification of Plant Lignin Content
 : FILE REFERENCE: 11000.1003c4U
 : CURRENT APPLICATION NUMBER: US/09/615.192A
 : CURRENT FILING DATE: 2000-07-12
 : PRIOR APPLICATION NUMBER: US 08/975.316
 : PRIOR FILING DATE: 1997-11-21
 : PRIOR APPLICATION NUMBER: US 08/713.000
 : PRIOR FILING DATE: 1996-09-11
 : PRIOR APPLICATION NUMBER: US 09/169.789
 : PRIOR FILING DATE: 1998-10-09

NUMBER OF SEQ ID NOS: 405
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 326
LENGTH: 114
TYPE: PR
ORGANISM: Eucalyptus grandis
US-09-615-192A-326

Query Match 11.9%: Score 434; DB 4; Length 114;
Best Local Similarity 71.1%: Pred. No. 1e-31;
Matches 81; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

QY 592 LNDHALNNGETKNTNITSPKATPEPEKYLKPEVGVIAVENDTSLPNKACR 651
DB 1 LNDQALVNESEINPSTSTFOKIVAFEEELKVLKVEVYVYETINLAIPLNKECR 60

QY 652 SYPLVYFVEBELGPGFTTGEKVTSPGEFTVFTAM*KGU110PLELVLAMNG 705
DB 61 SYPLVYKIVREHLOTALITGCGVITSPGDFDKVFTACAGKIDPLFCLSCWNG 114

RESULT 13
US-09-615-192A-329
Sequence 329, Application US/09615192A
Patent No. 6410718

GENERAL INFORMATION:
APPLICANT: Bloksberg, Leonard N.
TITLE OF INVENTION: Materials and Methods for the
TITLE OF INVENTION: Modification of Plant Lignin Content
FILE REFERENCE: 11000.1003c40
CURRENT APPLICATION NUMBER: US/09/615.192A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 08/975.316
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: US 08/713.000
PRIOR FILING DATE: 1996-09-11
PRIOR APPLICATION NUMBER: US 09/169.789
PRIOR FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 405
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 329
LENGTH: 97
TYPE: PR
ORGANISM: Eucalyptus grandis
US-09-615-192A-329

Query Match 10.7%: Score 389; DB 4; Length 97;
Best Local Similarity 76.3%: Pred. No. 3.3e-28;
Matches 74; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 182 NNNITPCLPLRTITASDVLPLSYIAGLLTGPNSKAVGPTGEVLNKAFAAGVEG 241
DB 1 NSGITPCLPLKOSISASGILVPPSYIAGLLTGPNSKAVGPTGEVLNKAFAAGVEG 60

QY 242 PFELOPKKELALVNCVAGSCMAVLPDANVALIUS 278
DB 61 PFELOPKKELALVNCVAGSCMAVLPDANVALIUS 97

RESULT 14
US-09-615-192A-331
Sequence 331, Application US/09615192A
Patent No. 6410718

GENERAL INFORMATION:
APPLICANT: Bloksberg, Leonard N.
APPLICANT: Havukkala, Ilkka
TITLE OF INVENTION: Materials and Methods for the
TITLE OF INVENTION: Modification of Plant Lignin Content
FILE REFERENCE: 11000.1003c40
CURRENT APPLICATION NUMBER: US/09/615.192A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 08/975.316

PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: US 08/713.000
PRIOR FILING DATE: 1996-09-11
PRIOR APPLICATION NUMBER: US 09/169.789
PRIOR FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 405
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 331
LENGTH: 132
TYPE: PR
ORGANISM: Pinus radiata
US-09-615-192A-331

Query Match 7.6%: Score 276; DB 4; Length 132;
Best Local Similarity 58.5%: Pred. No. 2.7e-17;
Matches 55; Conservative 13; Mismatches 24; Indels 2; Gaps 1;

QY 16 IKDPLMVCYAAALITGSHLDEKVKVAPKPVKLGCTITVSQVAGIAANSDIVKV 75
DB 41 LNDPLMMAAALSMGSHFEKVKRMHPRHRSIVVILQSGILIAQVAAVARIQS--VRV 98

QY 76 ELSEAARACVAKASDQVMSKNGTDSGVTTGF 109
DB 94 FLFETAKAVDPSSWVMMSMANGTDSGVTTGF 132

RESULT 15
US-09-268-347-36
Sequence 36, Application US/09268347
Patent No. 6335182

GENERAL INFORMATION:
APPLICANT: Icosmure, Shucua M.
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
FILE REFERENCE: 1038-860
CURRENT APPLICATION NUMBER: US/09/268.347
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 36
LENGTH: 2411
TYPE: PR
ORGANISM: Haemophilus influenzae
US-09-268-347-36

Query Match 3.4%: Score 122.5; DB 4; Length 2411;
Best Local Similarity 20.7%: Pred. No. 0.34;
Matches 144; Conservative 81; Mismatches 247; Indels 223; Gaps 33;

QY 2 ENGNHNVGVNNEICIKDPLNMGVA---AEALITGSHLDEKVKVAPKPVKLGCTIT 57
DB 347 ESDGDTTASVT---KIDNGSITVYDAKVGDKFSDDKIIVD---TTALITV 395

QY 58 VSDVAGIAA-----ANDSDIVKVELSEARAGVAAASDQVMSKNGTDSGVTT 107
DB 396 GGVAVARLAKRHDKKIKLVNAGDILVTALQNSWKAKAEDITDGALEGISKQGVKACQTV 455

QY 108 GFATSHRRTKQGA-----ANDSDIVKVELSEARAGVAAASDQVMSKNGTDSGVTT 136
DB 456 TFKAGKRLKVKQDGFANFTSLDALTGLTSLITVSGTIVGNDAKTVIKKDLITTPNG-- 513

QY 137 GNGTTSHTLPHSATRAAM-----VRIINTILQCYSGIRREIILAIATKPLNNITPCL 189
DB 514 NGTGTGTNTI--SVTKDKIKAGNKAIITVNSGLRAYDANDFDVYLNNSATDINRVEDAYK 571

QY 190 -----PLRGITIAS--GDLVPLSYIAGLLTGPNSKAVGPTGEVLNKAFAAGVEG 233
DB 572 GLINLEKKNANOPVATDSTAATVGDILKIKGVASTKNGIKFHSQVQAQAEV-----F 626

QY 234 AAGV-----EGGFEL-----QPKKIALV--KATA 258
DB 627 TGGAAATVTSKSENKHTITVSVAETKALUSLEKIGDITIKLVNDQNDVNLTVGNNKTA 686

QY 259 VQSCMSMVLIPDANVALIUSVIAISAFPAVMOCKPFTDHLITHKIKHHPGQIFAAAIMRY 318

Genome version 5.1.4.F5.457A
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 29, 2003, 02:12:21 : Search time 182 Seconds

(without alignments)
229.410 Million cell updates/sec

Title: US-09-964-992A-1

Perfect score: 1644

1 MENGINEVGVNELCKIDPL IDPLELTVWMEPLPT 711

Sequence: 1 MENGINEVGVNELCKIDPL

Scoring table: HUSUM62

Gapop 10.0, Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*\n2: /cgn2_6/ptodata/1/pubpaa/PRT_NEW_PUB pep:*\n3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*\n4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOM pep:*\n5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*\n6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOM pep:*\n7: /cgn2_6/ptodata/1/pubpaa/PCTIS_PUBCOM pep:*\n8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOM pep:*\n9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOM pep:*\n10: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*\n11: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOM pep:*\n12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOM pep:*\n13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOM pep:*\n14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOM pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3644	100.0	711	9	US-09-964-992A-1
2	3290.5	90.3	713	9	US-09-964-992A-2
3	3208	88.0	666	9	US-09-964-992A-5
4	2992	82.1	717	10	US-09-854-122-49
5	2728.5	74.9	700	10	US-09-854-122-50
6	1337	36.7	304	10	US-09-854-122-47
7	1332	36.6	303	10	US-09-854-122-48
8	1025	28.1	720	10	US-09-939-408A-13
9	994.5	27.3	716	10	US-09-765-873A-8
10	994.5	27.3	716	10	US-09-939-408A-19
11	991.5	27.2	716	10	US-09-765-873A-10
12	991.5	27.2	716	10	US-09-765-873A-34
13	990.5	27.2	716	10	US-09-765-873A-35
14	990.5	27.2	716	10	US-09-765-873A-37
15	988	27.1	686	10	US-09-765-873A-32
16	986.5	27.1	716	10	US-09-765-873A-38
17	984.5	27.0	716	10	US-09-765-873A-36
18	980.5	26.9	716	10	US-09-765-873A-33
19	971.5	26.7	713	10	US-09-939-408A-17

20	967	26.5	726	10	US-09-939-408A-21	Sequence 21, Appl
21	917	25.2	720	10	US-09-939-408A-30	Sequence 30, Appl
22	844.5	23.2	740	10	US-09-939-408A-15	Sequence 15, Appl
23	529	14.5	509	10	US-09-815-242-12085	Sequence 12085, A
24	527	14.5	513	10	US-09-833-745-50	Sequence 50, Appl
25	521.5	14.3	505	10	US-09-815-242-13970	Sequence 13970, A
26	518	14.2	513	10	US-09-833-745-51	Sequence 51, Appl
27	517	14.2	513	10	US-09-833-745-58	Sequence 58, Appl
28	513	14.1	509	10	US-09-833-745-34	Sequence 34, Appl
29	508	13.9	513	10	US-09-833-745-49	Sequence 49, Appl
30	505	13.9	508	10	US-09-833-745-40	Sequence 40, Appl
31	502	13.8	513	10	US-09-833-745-43	Sequence 43, Appl
32	501	13.7	513	10	US-09-833-745-53	Sequence 53, Appl
33	489	13.4	491	10	US-09-815-242-5671	Sequence 5671, Ap
34	489	13.4	504	10	US-09-815-242-12692	Sequence 12692, A
35	479	13.1	513	10	US-09-833-745-52	Sequence 52, Appl
36	476	13.1	504	10	US-09-833-745-95	Sequence 95, Appl
37	476	13.1	516	10	US-09-833-745-41	Sequence 41, Appl
38	473	13.0	513	10	US-09-833-745-46	Sequence 46, Appl
39	469.5	12.9	557	10	US-09-833-745-36	Sequence 36, Appl
40	468.5	12.9	513	10	US-09-833-745-92	Sequence 92, Appl
41	467.5	12.8	657	10	US-09-833-745-47	Sequence 47, Appl
42	461	12.7	513	10	US-09-833-745-48	Sequence 48, Appl
43	461	12.7	513	10	US-09-833-745-48	Sequence 48, Appl
44	458.5	12.6	513	10	US-09-833-745-45	Sequence 45, Appl
45	455.5	12.5	635	10	US-09-833-745-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-09-964-992A-1
Sequence 1, Application US/09964992A
Patient No. US2002073633A1
GENERAL INFORMATION:
APPLICANT: Sallveit, Mikal E.
APPLICANT: Campos, Reinaldo
APPLICANT: No. US2002073633A1logaki, Hiroyuki
APPLICANT: Sallveit, Mikal E.
APPLICANT: Sallveit, Mikal E.
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Characterization of Phenylalanine Ammonia Lyase (PAL)
FILE REFERENCE: 023070-124500ms
CURRENT APPLICATION NUMBER: US/09/964,992A
CURRENT FILING DATE: 2002-07-09
PRIOR APPLICATION NUMBER: US 60/245,956
PRIOR FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 711
TYPE: PRT
ORGANISM: Lactuca sativa
FEATURE:
OTHER INFORMATION: Lactuca phenylalanine ammonia lyase (PAL)
OTHER INFORMATION: (LSPAL)
US-09-964-992A-1

Query Match 100.0% Score 3644: DB 9: Length 711:
Post-local similarity 100.0% Prev No. 0: Gap 0:
Matches 711: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY 1 MENGINEVGVNELCKIDPLNMCVAAATGSHLDVKKMAVAFKKPVKLOGETITVSQ 60
QY 1 MENGINEVGVNELCKIDPLNMCVAAATGSHLDVKKMAVAFKKPVKLOGETITVSQ 60
DB 1 MENGINEVGVNELCKIDPLNMCVAAATGSHLDVKKMAVAFKKPVKLOGETITVSQ 60
QY 61 VAGIAAANDSTYKVELSEAKAGVAKSSDWMESNNKIDTSYCVTIGCATSHRRKQG 120
QY 61 VAGIAAANDSTYKVELSEAKAGVAKSSDWMESNNKIDTSYCVTIGCATSHRRKQG 120
DB 61 VAGIAAANDSTYKVELSEAKAGVAKSSDWMESNNKIDTSYCVTIGCATSHRRKQG 120
QY 121 GALKKELIFPLNNAIPENGTETSHLPHSATPAAMVPIINTLGVSGIFRELEATIKF 180
QY 121 GALKKELIFPLNNAIPENGTETSHLPHSATPAAMVPIINTLGVSGIFRELEATIKF 180

Query Match 88.0% Score 3208; DB 9; Length 666;
 Host Local Similarity 94.6% Pred No 176-238;
 Matches 628; Conservative 12; Mismatches 24; Indels 0; Gaps 0.

QY 1 MENCNNGVYNELCICDPLNMVAAEALTSGLDEYKKNVAEFPKPVYKLGSETLTVSQ 60
 DB 1 MENCNNGVYNELCICDPLNMVAAEALTSGLDEYKKNVAEFPKPVYKLGSETLTVSQ 60
 QY 61 VAGIAAADSTIVKVEHSEAKACVAKSSDWVMSNMKGCTDSYGVTTGPGATSHRRTKOC 120
 DB 61 VAGIAAADSTIVKVEHSEAKACVAKSSDWVMSNMKGCTDSYGVTTGPGATSHRRTKOC 120
 QY 121 GALOKELREFNAGIEFGTETSHTLPHSATRAAMVIRITLLQGYSGIFEELEATTKF 180
 DB 121 GALOKELREFNAGIEFGTETSHTLPHSATRAAMVIRITLLQGYSGIFEELEATTKF 180
 QY 181 LNNITPLPLRGITTTASGDIPLSYTAGLTPRPNKAVGPTGEVLNKAFAAAGVEG 240
 DB 181 LNNITPLPLRGITTTASGDIPLSYTAGLTPRPNKAVGPTGEVLNKAFAAAGVEG 240
 QY 241 GPFELQPEKGLAIVNGIAVSGMASMVLFDANVIALISVIAIFAEVMOGKPEFTDHLT 300
 DB 241 GPFELQPEKGLAIVNGIAVSGMASMVLFDANVIALISVIAIFAEVMOGKPEFTDHLT 300
 QY 301 HKLKHHPQVLEAAALMEYILDGSDYVKAQKVHEMDPLQKPKUUKYALRTSPQMLSPQLE 360
 DB 301 HKLKHHPQVLEAAALMEYILDGSDYVKAQKVHEMDPLQKPKUUKYALRTSPQMLSPQLE 360
 QY 361 VIRSTKMEIEINSVNDPLIDVSRKKAHLGSGNFQVITPGVSMONTRLAIAAGLMAFA 420
 DB 361 VIRSTKMEIEINSVNDPLIDVSRKKAHLGSGNFQVITPGVSMONTRLAIAAGLMAFA 420
 QY 421 QFSELVNFYNNGLPSNLSGGRNPILDYCKGCEIAMAASCSLQFPLANVITHVQSAEQ 480
 DB 421 QFSELVNFYNNGLPSNLSGGRNPILDYCKGCEIAMAASCSLQFPLANVITHVQSAEQ 480
 QY 481 HNOVNSIGLISAKKTAFAVDILKLMSTYVALCOSIDILHLENNKSTVKNVTSQVAK 540
 DB 481 HNOVNSIGLISAKKTAFAVDILKLMSTYVALCOSIDILHLENNKSTVKNVTSQVAK 540
 QY 541 KVLTVNGELHPSRFECKDLRVADREYFAYIDECVSTIPLMUKLRQVLDHALNNG 600
 DB 541 KVLTVNGELHPSRFECKDLRVADREYFAYIDECVSTIPLMUKLRQVLDHALNNG 600
 QY 601 ETEKNTNISIPOKIAFPEELKVLIPKRVGVRIAYENDTLISIPNRIKACRSYPIVHFVR 660
 DB 601 ETEKNTNISIPOKIAFPEELKVLIPKRVGVRIAYENDTLISIPNRIKACRSYPIVHFVR 660
 QY 661 DELG 664
 DB 661 DELG 664
 RESULT 4
 US-09-854-122-49
 : Sequence 49, Application US/09854122
 : Patent No. US20020016980A1
 : GENERAL INFORMATION:
 : APPLICANT: ALBERTE, RANDALL S.
 : APPLICANT: SMITH, ROBERT
 : TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
 : FILE REFERENCE: PHA-007-01
 : CURRENT APPLICATION NUMBER: US/09/854,122
 : CURRENT FILING DATE: 2001-09-10
 : PRIOR APPLICATION NUMBER: 60/202,529
 : PRIOR FILING DATE: 2000-05-10
 : NUMBER OF SEQ ID NOS: 51
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 49
 : TYPE: PRT
 : LENGTH: 717
 : ORGANISM: Arabidopsis thaliana

US-09-854-122-49

Query Match 82.1% Score 2922; DB 19; Length 717;
 Host Local Similarity 81.2% Pred No 7,96-222;
 Matches 575; Conservative 52; Mismatches 79; Indels 2; Gaps 1;

QY 4 GNVHGVYNELCICDPLNMVAAEALTSGLDEYKKNVAEFPKPVYKLGSETLTVSQVAG 64
 DB 12 GNVHGVYNELCICDPLNMVAAEALTSGLDEYKKNVAEFPKPVYKLGSETLTVSQVAG 64
 QY 64 1AANDSPITVVEHSEAKACVAKSSDWVMSNMKGCTDSYGVTTGPGATSHRRTKOCAL 123
 DB 64 1AANDSPITVVEHSEAKACVAKSSDWVMSNMKGCTDSYGVTTGPGATSHRRTKOCAL 123
 QY 124 OKELIRPLINAGIFGCTETSHTLPHSATRAAMVIRITLLQGYSGIFEELEATTKF 183
 DB 124 OKELIRPLINAGIFGCTETSHTLPHSATRAAMVIRITLLQGYSGIFEELEATTKF 183
 QY 184 NITPLPLRGITTTASGDIPLSYTAGLTPRPNKAVGPTGEVLNKAFAAAGVEGFEF 243
 DB 184 NITPLPLRGITTTASGDIPLSYTAGLTPRPNKAVGPTGEVLNKAFAAAGVEGFEF 243
 QY 244 ELQPEKGLAIVNGIAVSGMASMVLFDANVIALISVIAIFAEVMOGKPEFTDHLT 303
 DB 244 ELQPEKGLAIVNGIAVSGMASMVLFDANVIALISVIAIFAEVMOGKPEFTDHLT 303
 QY 304 KHHPOIFAAALMEYILDGSDYVKAQKVHEMDPLQKPKUUKYALRTSPQMLSPQLE 363
 DB 304 KHHPOIFAAALMEYILDGSDYVKAQKVHEMDPLQKPKUUKYALRTSPQMLSPQLE 363
 QY 364 SSKMEIEINSVNDPLIDVSRKKAHLGSGNFQVITPGVSMONTRLAIAAGLMAFA 423
 DB 364 SSKMEIEINSVNDPLIDVSRKKAHLGSGNFQVITPGVSMONTRLAIAAGLMAFA 423
 QY 424 ELVNFYNNGLPSNLSGGRNPILDYCKGCEIAMAASCSLQFPLANVITHVQSAEQ 483
 DB 424 ELVNFYNNGLPSNLSGGRNPILDYCKGCEIAMAASCSLQFPLANVITHVQSAEQ 483
 QY 484 DVNSIGLISAKKTAFAVDILKLMSTYVALCOSIDILHLENNKSTVKNVTSQVAK 543
 DB 484 DVNSIGLISAKKTAFAVDILKLMSTYVALCOSIDILHLENNKSTVKNVTSQVAK 543
 QY 544 TMGVNGELHPSRFECKDLRVADREYFAYIDECVSTIPLMUKLRQVLDHALNNGETE 603
 DB 544 TMGVNGELHPSRFECKDLRVADREYFAYIDECVSTIPLMUKLRQVLDHALNNGETE 603
 QY 604 KNTNISIPOKIAFPEELKVLIPKRVGVRIAYENDTLISIPNRIKACRSYPIVHFVR 663
 DB 604 KNTNISIPOKIAFPEELKVLIPKRVGVRIAYENDTLISIPNRIKACRSYPIVHFVR 663
 QY 664 GCGFLTGKKTSPGHEPDRVFIAACKQDILQPLAECGAGNNGEPLDPC 711
 DB 670 GCGFLTGKKTSPGHEPDRVFIAACKQDILQPLAECGAGNNGEPLDPC 717

RESULT 5

US-09-854-122-50
 : Sequence 50, Application US/09854122
 : Patent No. US20020016980A1
 : GENERAL INFORMATION:
 : APPLICANT: ALBERTE, RANDALL S.
 : APPLICANT: SMITH, ROBERT
 : TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
 : FILE REFERENCE: PHA-007-01
 : CURRENT APPLICATION NUMBER: US/09/854,122
 : CURRENT FILING DATE: 2001-09-10
 : PRIOR APPLICATION NUMBER: 60/202,529
 : PRIOR FILING DATE: 2000-05-10
 : NUMBER OF SEQ ID NOS: 51
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 50
 : TYPE: PRT
 : LENGTH: 700

ORGANISM: Trillium aestivum
US-09-854-122-50

Query Match: 74.9%; Score 2728.5; DB 10; Length 700;
Best Local Similarity 75.9%; Pred. No. 1.4e-201;
Matches 527; Conservative 96; Mismatches 96; Indels 5; Gaps 5;

QY 18 DPLMGAVALATGSHLDEKKAWEFFRRPVYVLTGETLTGVSQVAGIAAANDSDIVKEL 77
DB 12 DPLMGAVALATGSHLDEKKAWEFFRRPVYVLTGETLTGVSQVAGIAAANDSDIVKEL 77
QY 78 SEARAGVAKASVWVMSNMKGTDSTGYVTGGATSHRRTKQAGALOKELIFLNAGIFG 137
DB 68 DESARAGVAKASVWVMSNMKGTDSTGYVTGGATSHRRTKQAGALOKELIFLNAGIFG 127
QY 138 NGFETSHLPHSAIRAMIVRINTLLOGYSGRFELLETITPLNNITPCLPIRGTTTA 197
DB 128 TGID-GHVLPAATIRAMIVRINTLLOGYSGRFELLETITPLNNITPCLPIRGTTTA 186
QY 198 SGDLVPLSTYAGLTGRPNKAVGPTGEVLNKAFAAGVSGFELLPKRGALVNGT 257
DB 187 SGDLVPLSTYAGLTGRPNKAVGPTGEVLNKAFAAGVSGFELLPKRGALVNGT 246
QY 258 AVSGMASVPLFPAANTALLSEVLSAIFAEVNGKPEFTDHLTHKLIHPGQIEAAATME 317
DB 247 AVSGMASVPLFPAANTALLSEVLSAIFAEVNGKPEFTDHLTHKLIHPGQIEAAATME 306
QY 318 YLLDGSVYKAAQKVHMDP1QPKODRYALRTSPQWLGHOIEVIRHSTKMIEREINSYN 377
DB 307 HILEGSSYMLAKKLGELDP1MKPKODRYALRTSPQWLGHOIEVIRHSTKMIEREINSYN 366
QY 378 DNPLIDVSRKKALHGNFQGTPIGVSMOITRLAIAIGLMEFAQSELVNDFYNGDPSN 437
DB 367 DNPLIDVSRKKALHGNFQGTPIGVSMOITRLAIAIGLMEFAQSELVNDFYNGDPSN 426
QY 438 LSGGRNSLDYGRKGEIAAVASCSELOFLANPYTNHVOASQHOVDVNSLISARKTA 497
DB 427 LSGGRNSLDYGRKGEIAAVASCSELOFLANPYTNHVOASQHOVDVNSLISARKTA 486
QY 498 EAVIDILKMSSTYVALCOSIDIRHLENNKSTVKNFVSQVAKVITMVGNSLHPSRVC 557
DB 487 EAVIDILKMSSTYVALCOSIDIRHLENNKSTVKNFVSQVAKVITMVGNSLHPSRVC 546
QY 558 EXDLIRVYDREYFAYIDVCSSTYPLMKRLROYLVDAHLNNGETEKNTSIFOKIATF 617
DB 547 EXDLIRVYDREYFAYIDVCSSTYPLMKRLROYLVDAHLNNGETEKNTSIFOKIATF 606
QY 618 EELKVLLEPEVEGVRIAYENDLISIPNRIKACRSYPLRFVEELRGFLTGKCVTSFG 677
DB 607 EELKVLLEPEVEGVRIAYENDLISIPNRIKACRSYPLRFVEELRGFLTGKCVTSFG 666
QY 678 EEDRVATAMCQOIIDPLKLCGNGGEPPLIC 711
DB 667 EEDRVATAMCQOIIDPLKLCGNGGEPPLIC 700

RESULT 6
US-09-854-122-47

Sequence 47; Application US/09854122
Patent No. US20020016980A1
GENERAL INFORMATION:
APPLICANT: ALBERTE, RANDALL S.
TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
FILE REFERENCE: PHA-007.01
CURRENT APPLICATION NUMBER: US/09/854,122
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/202,529
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 47
LENGTH: 304

TYPE: PRT
ORGANISM: Zostera marina
US-09-854-122-47

Query Match: 36.7%; Score 1337; DB 10; Length 304;
Best Local Similarity 87.7%; Pred. No. 4.1e-95;
Matches 256; Conservative 20; Mismatches 16; Indels 0; Gaps 0;

QY 339 QKPKODRYALRTSPQWLGHOIEVIRHSTKMIEREINSYNVDNPLIDVSRKKALHGNFQGT 398
DB 13 KSKODRYALRTSPQWLGHOIEVIRHSTKMIEREINSYNVDNPLIDVSRKKALHGNFQGT 72
QY 399 PIGVSMOITRLAIAIGLMEFAQSELVNDFYNNGLPSNLSCGRNPSTDYGRKGEIAMA 458
DB 73 PIGVSMOITRLAIAIGLMEFAQSELVNDFYNNGLPSNLSCGRNPSTDYGRKGEIAMA 132
QY 459 SYCSELOFLANPYTNHVOASQHOVDVNSLGLISAKTAEAVDILKMSSTYVALCOSI 518
DB 133 SYCSELOFLANPYTNHVOASQHOVDVNSLGLISAKTAEAVDILKMSSTYVALCOSI 192
QY 519 DLKHEENKASVKNIVSOVAKVLMGVNGSLHPSRCEKDLKVVDRVYVFSTIDDPG 578
DB 193 DLKHEENKASVKNIVSOVAKVLMGVNGSLHPSRCEKDLKVVDRVYVFSTIDDPG 252
QY 579 SGTYPLMOKLRQVLDHALNNGETEKNTSIFOKIATFEEELKVLPEKVE 630
DB 253 SATYPLMOKLRQVLDHALNNGETEKNTSIFOKIATFEEELKVLPEKVE 304

RESULT 7
US-09-854-122-48

Sequence 48; Application US/09854122
Patent No. US20020016980A1
GENERAL INFORMATION:
APPLICANT: ALBERTE, RANDALL S.
TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
FILE REFERENCE: PHA-007.01
CURRENT APPLICATION NUMBER: US/09/854,122
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/202,529
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 48
LENGTH: 303
TYPE: PRT
ORGANISM: Zostera marina
US-09-854-122-48

Query Match: 36.6%; Score 1332; DB 10; Length 303;
Best Local Similarity 87.6%; Pred. No. 9.9e-95;
Matches 255; Conservative 20; Mismatches 16; Indels 0; Gaps 0;

QY 339 QKPKODRYALRTSPQWLGHOIEVIRHSTKMIEREINSYNVDNPLIDVSRKKALHGNFQGT 398
DB 13 KSKODRYALRTSPQWLGHOIEVIRHSTKMIEREINSYNVDNPLIDVSRKKALHGNFQGT 72
QY 399 PIGVSMOITRLAIAIGLMEFAQSELVNDFYNNGLPSNLSCGRNPSTDYGRKGEIAMA 458
DB 73 PIGVSMOITRLAIAIGLMEFAQSELVNDFYNNGLPSNLSCGRNPSTDYGRKGEIAMA 132
QY 459 SYCSELOFLANPYTNHVOASQHOVDVNSLGLISAKTAEAVDILKMSSTYVALCOSI 518
DB 133 SYCSELOFLANPYTNHVOASQHOVDVNSLGLISAKTAEAVDILKMSSTYVALCOSI 192
QY 519 DLKHEENKASVKNIVSOVAKVLMGVNGSLHPSRCEKDLKVVDRVYVFSTIDDPG 578
DB 193 DLKHEENKASVKNIVSOVAKVLMGVNGSLHPSRCEKDLKVVDRVYVFSTIDDPG 252
QY 579 SGTYPLMOKLRQVLDHALNNGETEKNTSIFOKIATFEEELKVLPEKVE 630
DB 253 SATYPLMOKLRQVLDHALNNGETEKNTSIFOKIATFEEELKVLPEKVE 303

RESULT 8

US-09-939-408A-13
 : Sequence 13, Application US/99/939408A
 : Patent No. US20020102712A1
 : GENERAL INFORMATION:
 : APPLICANT: Yoshida, Kōdōta
 : APPLICANT: Kōdōta, Anna
 : TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
 : TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
 : FILE REFERENCE: 29479/500NSCA
 : CURRENT APPLICATION NUMBER: US/09/939,408A
 : PRIOR FILING DATE: 2001-08-24
 : PRIOR APPLICATION NUMBER: US 09/624,693
 : PRIOR FILING DATE: 2000-07-24
 : PRIOR APPLICATION NUMBER: PCT/US01/23270
 : PRIOR FILING DATE: 2001-07-24
 : NUMBER OF SEQ ID NOS: 30
 : SOFTWARE: Patent Ver. 2.0
 : SEQ ID NO 13
 : LENGTH: 720
 : TYPE: PRT
 : ORGANISM: Rhodotorula glutinis
 : FEATURE:
 : NAME/KEY: unsure
 : LOCATION: (153)
 : OTHER INFORMATION: Other Information: Xaa - Val or Ala
 : US-09-939-408A-13

Query Match 28.18, Score 1025, DA 10, Length 720,
 Best Local Similarity 37.98, Pred. No. 1.5e-70;
 Matches 267, Conservative 127, Mismatches 259, Indels 52, Gaps 18,

QY 23 CVAALATGSHLDIVKKNVAFKKPVYKIGCEITITVSOVAGIAANDSTVYKELSHAR 82
 DB 40 GUCMAHQSLEIVQELSDPDDVVELSDYSLIKDVVG-AAKGRKVRQNDDELK 97
 QY 83 AGVAKSSWVESNMNKTDSYGTGATGSHRTKGGALOKELIFELNAGI----- 135
 DB 98 AKVAKSVIPLKAGUQNSV-YGVITGCGSADTKIELAVSLKALIRHQUGVPIRISXS 155
 QY 136 FONGTETSHITPHSATVYKAMIVIRINTLQOYSGIRPELLEATIKYNNITPCLPQGT 195
 DB 156 FSVGPLENTLPLEVPRAMVPIVNSLPGHSAYLVVLELTNELNHPITPIVLRGSI 215
 QY 196 TASGDLVPLSTIAGLITGRPNKSA---VGPTEGLVNAKKAPAAAGVCGPPELPKKGIA 252
 DB 216 SASGDLPLSTIAGLITGRPNKSA---VGPTEGLVNAKKAPAAAGVCGPPELPKKGIA 273
 QY 253 LVNGTAVGSGMAVLEFANVLLSEVLSAIEAVMGKRP-ETIDILTKLKHPRQTE 311
 DB 274 LVNGTAVGSGMAVLEFANVLLSEVLSAIEAVMGKRP-ETIDILTKLKHPRQTE 333
 QY 312 AAALIMETILDSQV-VAAQGVHHEMDPLQKQURVALRTSPQMLGPQIVIRISSTKME 370
 DB 334 VARNIRTLSSGSFAVEIEEYKVKDDEGILKQDRPLPTSPQGLPDELVDMMHVA 393
 QY 371 REINSVNDPLIDVSRNKALHCGNPGCPICVSMNDTRILAAIGKIMFQPSHINDPY 430
 DB 394 LEINTTTNPLIDVSRNKALHCGNPGCPICVSMNDTRILAAIGKIMFQPSHINDPY 453
 QY 431 NNGTSPNSGGRNPSLLVYFKGAEIEMASVYSELQFLANPTNIVVSAPVHNDVNSL 490
 DB 454 NNGTSPNSGGRNPSLLVYFKGAEIEMASVYSELQFLANPTNIVVSAPVHNDVNSL 512
 QY 491 ISAKTAAVNDILKIMSTYVALQSGIDILHLENNKSTVKNVSVQAKVITMG--VN 548
 DB 513 ISAKTAAVNDILKIMSTYVALQSGIDILHLENNKSTVKNVSVQAKVITMG--VN 572
 QY 549 G-ELHPSRFECKDLKLVUREY-----VFAYIDVSCSYPIPMOKIKOVIVDHAINGE 601
 DB 572 G-ELHPSRFECKDLKLVUREY-----VFAYIDVSCSYPIPMOKIKOVIVDHAINGE 617

DB 573 ALALFVKKALKRLKQITTYDLEPRHMDAFY-----ATGT-----VVELSSSP 617

QY 602 TEKNTNTSI-FOKIATFEELKVLIPKEVESVRAVENDTLSIPNLIKAKRSTPLREVR 660

DB 618 SANVTITAVNAMKVASAKKATS--LTREVPN-PEWJIPSSGAPAHVILSPRTVLSFVR 674

QY 661 FELG----RG-FITGEKVTSPGEFEDVFTAMCKGQIDPLILQCL 700

DB 675 BELGQVAPRGVFGVQDETIGSNVSRITYEATIKDGRINIVLVKML 719

RESULT 9

US-09-765-873A-8
 : Sequence 8, Application US/09765873A
 : Patent No. US20010053847A1
 : GENERAL INFORMATION:
 : APPLICANT: Tang, Xiao-Song
 : TITLE OF INVENTION: HOPRODUCTION OF HAA-HYDROXYCINNAMIC ACID
 : FILE REFERENCE: PCT009 US CIP
 : CURRENT APPLICATION NUMBER: US/09/765,873A
 : PRIOR FILING DATE: 2001-01-19
 : PRIOR APPLICATION NUMBER: US 09/627,216
 : PRIOR FILING DATE: 2000-07-27
 : PRIOR APPLICATION NUMBER: US 60/147,719
 : PRIOR FILING DATE: 1999-08-06
 : NUMBER OF SEQ ID NOS: 38
 : SOFTWARE: Microsoft Office 97
 : SEQ ID NO 8
 : LENGTH: 716
 : TYPE: PRT
 : ORGANISM: Rhodotorula glutinis
 : US-09-765-873A-8

Query Match 27.38, Score 994.5, DA 10, Length 716,
 Best Local Similarity 35.98, Pred. No. 3.2e-68;
 Matches 270, Conservative 123, Mismatches 250, Indels 109, Gaps 23;

QY 8 NGVNEELCIKIPIMMGVAAEALTGSHL-----DEKKVNAFPPKPVYKIGCEITITVS 59
 DB 14 NSVNS---AKAVNAGSNTNLAVNASHLPTTVTVVULVEKMLANPITDLELGYSLNLG 70
 QY 60 QVAGIAAANSTYKVEISFAAPAGVAKSSWVESNMNKTDSYGTGATGSHRTKGGAL 119
 DB 71 IV--VSAARKGPPRPVKQSDERISKIKSVETLPSQLSMV-YGVITGCGSADPRTFD 126
 QY 120 GSALQKELIFELNAGI-----FNGTETSHITPHSATVYKAMIVIRINTLQOYSGIR 170
 DB 127 ALSJQKALLHQUGCVLIPSSVDSRLQDGLNS--IDPLVVRGAMIRVNSLIRGSAVR 184
 QY 171 FELLKATIKPLNNITPQCTPQCTTFASGDI VPLSYIAGLITGRPNKSA-VGPTG--EVI 227
 DB 185 LVLEALTNPLNHTGTPVPLRGITISASGDLPSLYIAAISHIPESKIVHIEGKEKIL 244
 QY 228 NAEKRAAAGVCGPPELPKKGIALVNGTAVGSGMAVLEFANVLLSEVLSAIEAVMG 287
 DB 245 YAREAMALFENLEP--VVGPRKGLGVNGIAVSAAMATIALIHAMLSILSSSTLTMTYE 302
 QY 288 VMGQKP-ETIDILTKLKHPRQTEAAIMEDYILDSQV-VAAQGVHHEMDPLQKQUR 345
 DB 303 AMVGAGSFRHTPFDVTRPHRTQIVAGCNIRKILFGSKRPVHNRHKKVVDGCGILKQK 362
 QY 346 YALRTSPQMLGPQIVIRISSTKMIEREINSVNDPLIDVSRNKALHCGNPGCPICVSM 404
 DB 363 YPLTPSPQMLGPQIVIRISSTKMIEREINSVNDPLIDVSRNKALHCGNPGCPICVSM 422
 QY 405 DNTRILALAIKIMLFAQVSHLVNDFYNNNGIPSNISGGRNPSLLVYFKGAEIEMASV 464
 DB 423 ETRIGIAQIGKINFTQITTEMLNAGMNGILPSCIA-AHDPSLIRCKGLIDIAAATVSH 481
 QY 465 QFLANPTNIVVSAPVHNDVNSLISAKTAAVNDILKIMSTYVALQSGIDILHLENNK 524
 DB 482 GHLANPTVTHVQPAEMAGVANSIALISAKRTTISNIVLSILATHIYGVIAIDIKAIR 541

```
QY 525 ENKSTVKNVTSQVAKKVLTMGVNCELHPSRCEKDLRVVDREYFVAVIDVCSGYPL 584
DB 542 FEKK-----OFGPA-----IVSLIDQH-----GSAMT 565
QY 585 MOKRQVLDHALNNGETEKNTSIFOKIATFEELKVLPEYEGV-----632
DB 566 GSNLRDRIE-KVKNKLARLEQNTSYDLYPRWHAFPAQTVVEVLSLSLAAYNA 624
QY 633 -RIAYENDTLSPNRK-----ACRSP-----LYRVEBELG---RG--FL 668
DB 625 WKVAALSAISLTRQVRETFWMSASTSPALSPTSPTQILYAFVREBELGVAKARGDVF 684
QY 669 TGEKVTSPGEEDRVFTAMCKGQIIDPLECL 700
DB 685 GKQEVTT-GSNVSKTYEAIKSGRINVLKML 715

RESULT 10
US-09-939-408A-19
; Sequence 19, Application US/09939408A
; Patent No. US20020102712A1
; GENERAL INFORMATION:
; APPLICANT: Yoshida, Roberta
; TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
; TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
; FILE REFERENCE: 29479/500NSCA
; CURRENT APPLICATION NUMBER: US/09/939,408A
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: PCT/US01/23270
; PRIOR FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 19
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Rhodotorula toruloides
US-09-939-408A-19

Query Match 27.3%; Score 994.5; DB 10; Length 716;
Best Local Similarity 35.9%; Pred No. 3.2e-68;
Matches 270; Conservative 123; Mismatches 250; Indels 109; Gaps 23;

QY 8 NGVAVNELCICIDPLNMGVAAEALTGSHL-----DEVKKMVAEPRKPVVKGEGTITVS 59
DB 14 NGVAS---AKQAVNAGASTMLAVAGSHLPTTYQYQYDIYEKMLAAPDSTLELDGYSLNLG 70
QY 60 QVAGIAAANDSDTVKVELSEAAAGYKASSDWMWMSMKNKGTDSYGVTTGTCATSHKRTKQ 119
DB 71 DV--VSAARKGRPVVKDDEIRSKIDKSVEPLRSOLMSV--YGVTTGCGSADRTED 126
QY 120 GGLQKELREFLNAGI-----FGNGTETSHTLPHSATRAAMIVRINTLDGYSGIR 170
DB 127 AISLQKALLERHOLCGVLPSSFSFRIGRLENS--LPLEYVKGANTIKVNSLTRGHSAYR 184
QY 171 FEILEATKFLNNNTTPCLPRTGTTASGDLVPLSYIACLLTGRPNKA-VGPTG--EVL 227
DB 185 LVLEALTNELNIGITPIVPLRGTTISASGDLSPLSYIAAISGHPDSKVHVHNEKEKIL 244
QY 228 NAKKATPAAGVGEFPELOPKEGIALVNGTAVSGMASMVLPDANVALISVLSAIFAE 287
DB 245 YARAAALFNLEP--VVLGPKRGIGLVNGTAVSASMATIALHDAMLSLSSQSLTAMTVE 302
QY 288 VMQGRP-EFTDHLTHKLRKLNHPQIEAAAIMLEYILDGSDY-VKAAQKVHEMDPLQPKQDR 345
DB 303 AMVGHAGSFPRFLDYTRPHPTQIEVAGNIKKLJCSRRVAHNEBKVUKDQDGLTRDR 362
QY 346 YALRTSPQWLGPOI-EVINSSTKMIEREJNSVNDNPLDIVSRNKALHCGNFGPTIGVSM 404
DB 363 YPLRTSPQWLGPLVDLIIHAHVLITEAGOSTDNDLLIDVENKTSHHGNGFOAAVANTM 422
```

```
QY 405 DNTPLAIAIKLMPAOPSELVNDVFYNNGLPSNLGSPNPSLDYGGKGETAMASGSEL 464
DB 423 EKRTRGLAQIGKLNFTQULEMLNAGNNRQLEPCLA-AEDPSLSYHCKQLDIAAAVISEL 481
QY 465 QPLANPVNTHVCSAEQNDQVNSLGLISARKTAEAVDLIKMSSTYVALCOSIDLRHLE 524
DB 482 GHILANPVTHVCPAEAMANOAVNSLALISARPTESNDVLSLLATHCLCYQVQALDLBAIE 541
QY 525 ENKSTVKNVTSQVAKKVLTMGVNCELHPSRCEKDLRVVDREYFVAVIDVCSGYPL 584
DB 542 FEKK-----OFGPA-----IVSLIDQH-----GSAMT 565
QY 585 MOKRQVLDHALNNGETEKNTSIFOKIATFEELKVLPEYEGV-----632
DB 566 GSNLRDRIE-KVKNKLARLEQNTSYDLYPRWHAFPAQTVVEVLSLSLAAYNA 624
QY 633 -RIAYENDTLSPNRK-----ACRSP-----LYRVEBELG---RG--FL 668
DB 625 WKVAALSAISLTRQVRETFWMSASTSPALSPTSPTQILYAFVREBELGVAKARGDVF 684
QY 669 TGEKVTSPGEEDRVFTAMCKGQIIDPLECL 700
DB 685 GKQEVTT-GSNVSKTYEAIKSGRINVLKML 715

RESULT 11
US-09-765-873A-10
; Sequence 10, Application US/09765873A
; Patent No. US20010053847A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Xiao-Song
; TITLE OF INVENTION: BIOPRODUCTION OF PAKA-HYDROXYCINNAMIC ACID
; FILE REFERENCE: BC1009 US CIP
; CURRENT APPLICATION NUMBER: US/09/765,873A
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 09/627,216
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/147,719
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 716
; TYPE: PRT
; ORGANISM: mutant from Rhodotorula glutinis
US-09-765-873A-10

Query Match 27.2%; Score 991.5; DB 10; Length 716;
Best Local Similarity 35.9%; Pred No. 5.5e-68;
Matches 270; Conservative 122; Mismatches 251; Indels 109; Gaps 23;

QY 8 NGVAVNELCICIDPLNMGVAAEALTGSHL-----DEVKKMVAEPRKPVVKGEGTITVS 59
DB 14 NGVAS---AKQAVNAGASTMLAVAGSHLPTTYQYQYDIYEKMLAAPDSTLELDGYSLNLG 70
QY 60 QVAGIAAANDSDTVKVELSEAAAGYKASSDWMWMSMKNKGTDSYGVTTGTCATSHKRTKQ 119
DB 71 DV--VSAARKGRPVVKDDEIRSKIDKSVEPLRSOLMSV--YGVTTGCGSADRTED 126
QY 120 GGLQKELREFLNAGI-----FGNGTETSHTLPHSATRAAMIVRINTLDGYSGIR 170
DB 127 AISLQKALLERHOLCGVLPSSFSFRIGRLENS--LPLEYVKGANTIKVNSLTRGHSAYR 184
QY 171 FEILEATKFLNNNTTPCLPRTGTTASGDLVPLSYIACLLTGRPNKA-VGPTG--EVL 227
DB 185 LVLEALTNELNIGITPIVPLRGTTISASGDLSPLSYIAAISGHPDSKVHVHNEKEKIL 244
QY 228 NAKKATPAAGVGEFPELOPKEGIALVNGTAVSGMASMVLPDANVALISVLSAIFAE 287
DB 245 YARAAALFNLEP--VVLGPKRGIGLVNGTAVSASMATIALHDAMLSLSSQSLTAMTVE 302
QY 288 VMQGRP-EFTDHLTHKLRKLNHPQIEAAAIMLEYILDGSDY-VKAAQKVHEMDPLQPKQDR 345
```

```

Db 303 AMVHGAGSFHPLHDVTPRPTDTEVAGNIPKLESGSPFAVHHEEVEVKKDEBGIILPQDP 362
QY 346 YALRTSOWMIGPOL-EVIRNSSTKIMREINSVNDPILIVSNKALHCGNPGIPIGVSM 404
Db 363 YPLRTSPWMLGPIVLSDLIHAHVLITTEAGOSTTDNPLINVENKTSIHGKGFQAAVANTM 422
QY 405 DNTFLIAALGKLMFAQFSELVNDFYNNGLSPNLSGGRPSLDYFPGSGELTMASSGSEL 464
Db 423 EKRIRIGIAOIGKLNFTQITPMIACGNKQIPSCIA-ARDPISLSTHCKGIDIAAAAVTSEL 481
QY 465 QFLANPVTHNVQARODVNSIGLISAKRTAEAVDILKIMSTYVALVGSIDLEHLE 524
Db 482 GHILANPVTHIVQAEKMANOVNSLALISARPTESNDVCSLLATHLVGVCAIDLEATE 541
QY 525 ENKSTVKNTVSQAKKVLITMGVNGHLHPSRCEDLKVVPREVPFVAVITPVGSGTPT 584
Db 542 FEFRK-----QFGPA-----IVSLIDQH-----GSAMT 565
QY 585 MUKRLVYLHIALNNGETEKNNTSIFOKIATFEELKYLLEKEGV-----632
Db 566 GSNLRDELVE-KVNKTLAKRLHQTNSYDILVPRHDAFS-AAGTVFVVISITSLAAVNA 624
QY 633 -RIAYENDTISINRIK-----ACRSY-----LYRVRRELG-----RG--FL 668
Db 625 WKVAAMESAISLTROYRETFWSMASTSPALSYLSEFTQILYAFVFEELGVKARRGDVFL 684
QY 669 TGEKVTSPGSEDFRVTAMCKGQIIDPLECL 700
Db 685 GKQEVIT-GSNVSKITYEAIKSGHINNVLLKML 715

RESULT 12
US-09-765-873A-34
Sequence 34, Application US/09765873A
Patent No. US20010053847A1
GENERAL INFORMATION:
APPLICANT: Tang, Xiao-Song
TITLE OF INVENTION: HIOPRODUCTION OF PAPA-HYDROXYCINNAMIC ACID
FILE REFERENCE: H01009 US CIP
CURRENT APPLICATION NUMBER: US/09/765,873A
CURRENT FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 09/627,216
PRIOR FILING DATE: 2000-07-27
PRIOR APPLICATION NUMBER: US 60/147,719
PRIOR FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 34
LENGTH: 716
TYPE: PRT
ORGANISM: mutant from Rhodotorula glutinis
FEATURE:
NAME/KEY: UNSURE
LOCATION: (126) (126)
OTHER INFORMATION: X= Gly, Ala, Ser, Thr
NAME/KEY: UNSURE
LOCATION: (138) (138)
OTHER INFORMATION: X= Leu, Met, Ile, Val, Cys
NAME/KEY: UNSURE
LOCATION: (540) (540)
OTHER INFORMATION: X= Thr, Ala, Ser, Pro, Gly
US-09-765-873A-34

Query Match 27-28, Score 991.5, Dh 10, Length 716,
Best Local Similarity 35.9%; Pred. No. 5,5e-68;
Matches 270; Conservative 122; Mismatches 251; Indels 109; Gaps 23;
QY 8 NGVNEHCIDPIIMGVAAALITGSH-----DEYKMKMAFERKPPVYKIGCTITVS 59
Db 14 NGVAA--AAQAVNCASTINAVAGSHPTTQVTVYVIVKMKLAAPIDSLIDLDVYSTING 70
QY 60 QVAGIAAANDSTVKEVLEAFARAGVAKASDWMESMNGTDSYGVTTSGFATSHIRIK 119

```

```

Db 71 DV--VSAAPKGPVPVPRKNDPILPSKIDKSVFELPSGLNSV--YGVITGTPYSATRETEX 126
QY 120 GGAUOKHLIKHINAGT-----PGNGTESHTIIPHSATIAAMIVHINILIDVYSIR 170
Db 127 ATSLQKALLEHLCGLVLPSSPDSFPLSPGLENS--LPELVYGMATIVNSLITKHSAYR 184
QY 171 FEILENITKRLNNNTPTPLPRTTASGDPLPSYIACTILGPRNSGA-VAPPTG--EVL 227
Db 185 LVVLEALTNINLNGITPPIVPLKCTISASGDSLPIYIAAALISGHIDSKSVHVVHREKKEIL 244
QY 228 NAEKAEAAVCEGCFEELQPKREGIALVNGTAAVGSAGSAVLEDAVVALLESEVSAIFAE 287
Db 245 YAREMAALEFLEP--VVLGKREDELGVNCTINVASMAIILALHAILMLLSGLSTAMTVE 302
QY 288 VPOCKP-EPTDHTLHKIKHPPQIEAAAIIMEYIIDCSIV--VKAACKVHMDHIVQIKQOR 345
Db 303 AMVHGAGSFHPLHDVTPRPTDTEVAGNIPKLESGSPFAVHHEEVEVKKDEBGIILPQDP 362
QY 346 YALRTSOWMIGPOL-EVIRNSSTKIMREINSVNDPILIVSNKALHCGNPGIPIGVSM 404
Db 363 YPLRTSPWMLGPIVLSDLIHAHVLITTEAGOSTTDNPLINVENKTSIHGKGFQAAVANTM 422
QY 405 DNTFLIAALGKLMFAQFSELVNDFYNNGLSPNLSGGRPSLDYFPGSGELTMASSGSEL 464
Db 423 EKRIRIGIAOIGKLNFTQITPMIACGNKQIPSCIA-ARDPISLSTHCKGIDIAAAAVTSEL 481
QY 465 QFLANPVTHNVQARODVNSIGLISAKRTAEAVDILKIMSTYVALVGSIDLEHLE 524
Db 482 GHILANPVTHIVQAEKMANOVNSLALISARPTESNDVCSLLATHLVGVCAIDLEATE 541
QY 525 ENKSTVKNTVSQAKKVLITMGVNGHLHPSRCEDLKVVPREVPFVAVITPVGSGTPT 584
Db 542 FEFRK-----QFGPA-----IVSLIDQH-----GSAMT 565
QY 585 MUKRLVYLHIALNNGETEKNNTSIFOKIATFEELKYLLEKEGV-----632
Db 566 GSNLRDELVE-KVNKTLAKRLHQTNSYDILVPRHDAFS-AAGTVFVVISITSLAAVNA 624
QY 633 -RIAYENDTISINRIK-----ACRSY-----LYRVRRELG-----RG--FL 668
Db 625 WKVAAMESAISLTROYRETFWSMASTSPALSYLSEFTQILYAFVFEELGVKARRGDVFL 684
QY 669 TGEKVTSPGSEDFRVTAMCKGQIIDPLECL 700
Db 685 GKQEVIT-GSNVSKITYEAIKSGHINNVLLKML 715

RESULT 13
US-09-765-873A-35
Sequence 35, Application US/09765873A
Patent No. US20010053847A1
GENERAL INFORMATION:
APPLICANT: Tang, Xiao-Song
TITLE OF INVENTION: BIOPRODUCTION OF PAPA-HYDROXYCINNAMIC ACID
FILE REFERENCE: H01009 US CIP
CURRENT APPLICATION NUMBER: US/09/765,873A
CURRENT FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 09/627,216
PRIOR FILING DATE: 2000-07-27
PRIOR APPLICATION NUMBER: US 60/147,719
PRIOR FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 35
LENGTH: 716
TYPE: PRT
ORGANISM: mutant from Rhodotorula glutinis
FEATURE:
NAME/KEY: UNSURE
LOCATION: (198) (198)
OTHER INFORMATION: X= Asp, Asn, Glu, Gln
NAME/KEY: UNSURE

```

LOCATION: (540) (540)
 OTHER INFORMATION: X= Thr, Ala, Ser, Pro, Gly
 US-09-765-873A-35

Query Match 27.2% Score 990.5; DB 10; Length 716;
 Best Local Similarity 35.9%; Pred No. 6, 6e-68;
 Matches 270; Conservative 122; Mismatches 251; Indels 109; Gaps 23;

```

QY 8 NGVNEICIKDPLKMGVAFAALTSGL-----DEYKKVAEFKKYVKLGSETLTVS 59
DB 14 NGVAS---AKQAVNGASTINLAVAGSHLPTTOYVQVDIVKMLAAPDSTLELDGYSILNG 70
QY 60 QVAGIAANDSDYKVELSEARAGVAKASDPMVSNKKGTDSYVTTGATSHHRTKO 119
DB 71 DY--VSAARKGRPVAKDSDEIRSKIDKSVFELRSQLSMSV--YGVTTGGASDRTED 126
QY 120 GGAIOKELIRPLNGI-----FCNGTSHLPHSATRAMIVRINTLIQYSGIR 170
DB 127 AISTOKALLEIOLCGVLPSSDFSLRGLGLEN--LPLEVRCAMTIRVNSLTRGHSAVR 184
QY 171 FELIATIKFLNNNTPTPLRGTTASGDLVPLSYAGLITGRPNKA--VGPTG--EVL 227
DB 185 LVLALNPLNHLXITPLVPLKGTISASGDLSPSTIAAALISGHPDKVHVHKGKEL 244
QY 228 NAEKAPAAAGVEGFEELQPEGLALNGTAVSGMASVLPDANVLLSEVLALFAE 287
DB 245 YAREMALFNLER--VVLGPKRGGLVNGTAVSASMATLALHDAHMLSLSSLTAMTVE 302
QY 288 VMQGRP--EFTDHLTKLKHHPGQIEAAALMEYILDGSDY--VKAQKVHEMDPIQKPRDR 345
DB 303 AMVAGAGSHPLHDVTRPHPTQIEVAGNIKLEGSFAVHHEEVKVKVDDDEGILRODR 362
QY 346 VALRTPQWMLGPOI--EVIKRSSTKMLEREINSVNDPPLIDVSRNKALHGNPQGTPIGVSM 404
DB 363 YPLRTSPQWMLGPOI--EVIKRSSTKMLEREINSVNDPPLIDVSRNKALHGNPQGTPIGVSM 422
QY 405 DNTLALIAIKLMAFQPSLELVNDYNNGLPSNLSSGNRPDLQSGFSGSELAAMASYSEL 464
DB 423 EKTRGLAOIGKLNTOLEMLNAGMNRGLPSCLA--AEDESLSHYCKGLDIAAAAYTSEL 481
QY 465 QFLANPVTNNHQSADQHNQDVNSLGLISARKTAAVAVDLLKMSSTYVALCOSIDLRLHE 524
DB 482 GHLANPVTTHQAPAMAAQAVNSLALISARKTESNVLSLLATHLXYCVLQALIDIRAXE 541
QY 525 ENMKSTVKNVTSQVAKKVLTMGVNGLHPSRFCEKDLIRVYDRKYVAYIDVCSSTYPL 564
DB 542 PEFKK-----QFGPA-----IVSLIDOHF-----GSAMT 565
QY 585 MOKLRQVLVDHALNNGTEKNTNTSIFQKIATFEEELKVLLEPKVEGV-----632
DB 566 GSNLDELVE--KVNKTIAKRLKQNTSYDLVPRMHDASFAGTYVEVLSTSLSLAAVNA 624
QY 633 -RIAYENDTLSPNRIK-----ACRSYP-----LYPRVRELG---KG--FL 668
DB 625 WKVAAESAISLITROVRETFWASASTSSPALSYSPRTQILYAFVREELGVAKKRDVFL 684
QY 669 TGEKVTSPGEEDRVFTAMCKGQIIDLLECL 700
DB 685 GKQEVYI--GSNVSKTYEAIKSGRIINVLKML 715

```

RESULT 14

US-09-765-873A-37
 Sequence 37, Application US/09765873A
 Patent No. US20010053847A1
 GENERAL INFORMATION:
 APPLICANT: Tang, Xiao-Song
 TITLE OF INVENTION: BIOPRODUCTION OF PARA-HYDROXYCINNAMIC ACID
 FILE REFERENCE: H01009 US CIP
 CURRENT APPLICATION NUMBER: US/09/765,873A
 PRIOR APPLICATION NUMBER: US 09/627,216
 PRIOR FILING DATE: 2000-07-27

PRIOR APPLICATION NUMBER: US 60/147,719
 PRIOR FILING DATE: 1999-08-06
 NUMBER OF SEQ. ID NOS: 38
 SOFTWARE: Microsoft Office 97
 SEQ. ID NO 37
 LENGTH: 716
 TYPE: PRT
 ORGANISM: mutant from Rhodotorula glutinis
 FEATURE:
 NAME/KEY: UNSURE
 LOCATION: (149)..(149)
 OTHER INFORMATION: X= Pro, Ala, Ser, Thr, Gly
 NAME/KEY: UNSURE
 LOCATION: (202)..(202)
 OTHER INFORMATION: X= Val, Met, Leu, Cys
 NAME/KEY: UNSURE
 LOCATION: (540)..(540)
 OTHER INFORMATION: X= Thr, Ala, Ser, Pro, Gly
 US-09-765-873A-37

Query Match 27.2% Score 990.5; DB 10; Length 716;
 Best Local Similarity 35.9%; Pred No. 6, 6e-68;
 Matches 270; Conservative 122; Mismatches 251; Indels 109; Gaps 23;

```

QY 8 NGVNEICIKDPLKMGVAFAALTSGL-----DEYKKVAEFKKYVKLGSETLTVS 59
DB 14 NGVAS---AKQAVNGASTINLAVAGSHLPTTOYVQVDIVKMLAAPDSTLELDGYSILNG 70
QY 60 QVAGIAANDSDYKVELSEARAGVAKASDPMVSNKKGTDSYVTTGATSHHRTKO 119
DB 71 DY--VSAARKGRPVAKDSDEIRSKIDKSVFELRSQLSMSV--YGVTTGGASDRTED 126
QY 120 GGAIOKELIRPLNGI-----FCNGTSHLPHSATRAMIVRINTLIQYSGIR 170
DB 127 AISTOKALLEIOLCGVLPSSDFSLRGLGLEN--LPLEVRCAMTIRVNSLTRGHSAVR 184
QY 171 FELIATIKFLNNNTPTPLRGTTASGDLVPLSYAGLITGRPNKA--VGPTG--EVL 227
DB 185 LVLALNPLNHLXITPLVPLKGTISASGDLSPSTIAAALISGHPDKVHVHKGKEL 244
QY 228 NAEKAPAAAGVEGFEELQPEGLALNGTAVSGMASVLPDANVLLSEVLALFAE 287
DB 245 YAREMALFNLER--VVLGPKRGGLVNGTAVSASMATLALHDAHMLSLSSLTAMTVE 302
QY 288 VMQGRP--EFTDHLTKLKHHPGQIEAAALMEYILDGSDY--VKAQKVHEMDPIQKPRDR 345
DB 303 AMVAGAGSHPLHDVTRPHPTQIEVAGNIKLEGSFAVHHEEVKVKVDDDEGILRODR 362
QY 346 VALRTPQWMLGPOI--EVIKRSSTKMLEREINSVNDPPLIDVSRNKALHGNPQGTPIGVSM 404
DB 363 YPLRTSPQWMLGPOI--EVIKRSSTKMLEREINSVNDPPLIDVSRNKALHGNPQGTPIGVSM 422
QY 405 DNTLALIAIKLMAFQPSLELVNDYNNGLPSNLSSGNRPDLQSGFSGSELAAMASYSEL 464
DB 423 EKTRGLAOIGKLNTOLEMLNAGMNRGLPSCLA--AEDESLSHYCKGLDIAAAAYTSEL 481
QY 465 QFLANPVTNNHQSADQHNQDVNSLGLISARKTAAVAVDLLKMSSTYVALCOSIDLRLHE 524
DB 482 GHLANPVTTHQAPAMAAQAVNSLALISARKTESNVLSLLATHLXYCVLQALIDIRAXE 541
QY 525 ENMKSTVKNVTSQVAKKVLTMGVNGLHPSRFCEKDLIRVYDRKYVAYIDVCSSTYPL 584
DB 542 PEFKK-----QFGPA-----IVSLIDOHF-----GSAMT 565
QY 585 MOKLRQVLVDHALNNGTEKNTNTSIFQKIATFEEELKVLLEPKVEGV-----632
DB 566 GSNLDELVE--KVNKTIAKRLKQNTSYDLVPRMHDASFAGTYVEVLSTSLSLAAVNA 624
QY 633 -RIAYENDTLSPNRIK-----ACRSYP-----LYPRVRELG---KG--FL 668
DB 625 WKVAAESAISLITROVRETFWASASTSSPALSYSPRTQILYAFVREELGVAKKRDVFL 684
QY 669 TGEKVTSPGEEDRVFTAMCKGQIIDLLECL 700

```

Db 685 GQGVTTI-GSNVSKTYEALIKSGRIINNVLKML 715

RESULT 15

US-09-765-873A-32
 ? Sequence 32, Application US/09765873A
 ? Patent No. US20010053847A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Tansu, Xiao-Song
 ? TITLE OF INVENTION: BIOPROTECTION OF PAPA-HITPO-XY-INNAMIN Acid
 ? FILE REFERENCE: H01009 US CIP
 ? CURRENT APPLICATION NUMBER: US/09/765,873A
 ? CURRENT FILING DATE: 2001-01-19
 ? PRIOR APPLICATION NUMBER: US 09/627,216
 ? PRIOR FILING DATE: 2000-07-27
 ? PRIOR APPLICATION NUMBER: US 60/147,719
 ? PRIOR FILING DATE: 1999-08-06
 ? NUMBER OF SEQ ID NOS: 38
 ? SOFTWARE: Microsoft Office 97
 ? SEQ ID NO 32
 ? LENGTH: 686
 ? TYPE: PRT
 ? ORGANISM: mutant from Rhodotorula glutinis
 ? US-09-765-873A-32

Query Match 27.1% Score 988; DB 10; Length 686;

Best Local Similarity 36.2%; Pred No. 9, 6e-68;

Matches 265; Conservative 121; Mismatches 240; Indels 106; Gaps 22.

QY 28 ALTGSHL-----DEYKKVAFPKIVVKKLGGFETLVSVACIAAANDPIVKKELSE 79
 Db 1 AVAGSHLPTQVTVYDIYERKMLAAPDSTLELDGYSINLGDV--VSAARKGRPYRVKSD 58
 QY 80 ARAQVYKASSDMYMSNMKKTDSYGVTTGPGATSHPTKQSGALQKELIFLNAGT----- 135
 Db 59 EIRSKIDKSVFPIKRSQISMV--YGVTTGPGSADPTREDAISLOKALIRHQLCGVLPSS 116
 QY 136 -----FGMGTETSIPLSATRAAIVRINTLQYSQIFELIETKFLNNITPLP 190
 Db 117 FDSFRLGGLENS--LPLEVVRGAMTIEVNSLTGRHSVRLVLEALTNFLNGITPIY 174
 QY 191 LMGTTTASGDUVPISYAGLITGRPNKA--VGPFG--FVINAERKAFACVGGCFERLQ 247
 Db 175 LRGTTASGSDPLSYIAAAGISHPDSKVHVHGEKKEIYAREAMALFNLEP--VVLGP 232
 QY 248 KEGLLVNGTVAGSGMASVLEFDANVLALLSEVLATFAEVMQGR--EFTDHLTHKLNH 306
 Db 233 KRGIGLVNGTAVASAMATLALHDAHMLSLISOSTAMTVEAMVGHAGSFHPPLHDVTRPH 292
 QY 307 PQGLEMAAMEYILDGSDY--VKAAGKVHEMDPLVKPKDRLALKTSFQWLGPOI-EVIRS 364
 Db 293 PTQIEVAGNIRKLEGSREFAVHHEEYVKVDDSGILRDYRPLRTPSPQWLGPLVSDLIHA 352
 QY 365 STKMLERKINSVNDPILIDVSKNKALHGNVGTPIGVSMNTRKLAIAAIGKLMFAQFS 424
 Db 353 HAVLTIEAGQSTDPNPLIDVENKTSHHGNGFOAAAVANTMEKTRLGIAQIGKLNFTOLTE 412
 QY 425 LVNDFYNNGLSPNSLGSFNPSTLFGFKGGEIYMASSCELOFLANPVTNIVUSAEQHNOD 484
 Db 413 MNAQMNNGIPLSCIA-ARDPISLHYCKKIDIAAAAYTSHGLANVPVTHVQPAEMANQA 471
 QY 485 VNSLGLISARKTAENVDLTKLMSSTYVALGQSIDLRHLEENKSTVKNTVSQYAKKVL 544
 Db 472 VNSLGLISARKTTSNDVSLTLATHLYCVLAIDLVAIEFEFK----- 516
 QY 545 MGVNCHLHNSRCEKDLIRVDREYFAYIDDVSGTYPIIMOKIROVIVDAHANNCHTER 604
 Db 517 -----QFGPA-----IVSLIDQHF-----GSAMTGSNLRDELVE-KVAKTLAKR 554
 QY 605 NNTSTIFOKIAIEEELKVLPLKEVEGV-----RIAYENDTSLIPNRK--- 648
 Db 555 LKQTSYDILVPRWHDAFSAAGIVVEVILSTSLIAAVNAWKVAASASISITPROVRETF 614

QY 649 ---ACRSYP-----LYRFRREIG-----KG--PIIGKQVTSPOGPHURVFTAMC 688
 Db 615 WSAASTSSPALSYLSPPTGTLVAFVPEHIAVKAPRGNVPIGKGEVTTI-GSNVSKTYEALIR 673
 QY 689 KQGLIDPLLECI 700
 Db 674 SGRIINNVLKML 685

Search completed: March 29, 2003, 02:54:43
 Job time: 184 secs

GenCore version 5.1.4 P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

0M protein - protein search, using sw model

Run on: March 29, 2003, 02:05:46 : Search time 49 seconds
(without alignments)
1394.930 Million cell updates/sec

Title: US-09-964-992A-1
Perfect score: 3644
Sequence: 1 MENNHVNCVYVNEI/TKDPL 11DPLECT/GWNGEPLPIC 711

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_73:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	3206	88.0	667	T12749	phenylalanine amno
2	3186.5	87.4	716	S04463	phenylalanine amno
3	3186.5	87.4	718	S48726	phenylalanine amno
4	3185.5	87.4	716	S48725	phenylalanine amno
5	3171.5	87.0	708	T14295	phenylalanine amno
6	3149	86.4	715	S66343	phenylalanine amno
7	3118.5	85.6	712	T03663	phenylalanine amno
8	3116	85.5	715	J02265	phenylalanine amno
9	3109.5	85.3	712	T01858	phenylalanine amno
10	3077.5	84.5	710	J05872	phenylalanine amno
11	3070	84.2	723	S25901	phenylalanine amno
12	3068	84.2	713	S22991	phenylalanine amno
13	3067.5	84.2	721	A44133	phenylalanine amno
14	3067.5	84.2	725	S17444	phenylalanine amno
15	3058	83.9	712	S04127	phenylalanine amno
16	3051	83.7	705	J05873	phenylalanine amno
17	3049	83.7	720	S21174	phenylalanine amno
18	3045.5	83.6	725	S64787	phenylalanine amno
19	3039.5	83.4	725	S52990	phenylalanine amno
20	3004	82.4	717	T146172	phenylalanine amno
21	2988	82.0	708	T109397	phenylalanine amno
22	2985.5	81.9	710	S60042	phenylalanine amno
23	2953.5	81.1	707	S29029	phenylalanine amno
24	2730.5	74.9	700	T06545	phenylalanine amno
25	2671	73.3	619	S56312	phenylalanine amno
26	2632.5	72.2	710	S66313	phenylalanine amno
27	2598	71.3	582	S60043	phenylalanine amno
28	2590	71.1	710	S04128	phenylalanine amno
29	2578.5	70.8	695	S52992	phenylalanine amno

30	2555.5	70.1	716	S28185	phenylalanine amno
31	2538	69.6	590	S70916	phenylalanine amno
32	2496	68.5	701	S06475	phenylalanine amno
33	2373.5	65.1	754	T09777	phenylalanine amno
34	2220	60.9	505	A24727	phenylalanine amno
35	2220	60.9	506	S04129	phenylalanine amno
36	2184	59.4	549	T05970	phenylalanine amno
37	1948	53.5	475	T05966	phenylalanine amno
38	1887	51.8	497	T05968	phenylalanine amno
39	1867	51.2	416	J01070	phenylalanine amno
40	1368	37.5	339	T06546	phenylalanine amno
41	994.5	27.3	716	A56628	phenylalanine amno
42	979	26.9	235	S25538	phenylalanine amno
43	977	26.8	693	A29607	phenylalanine amno
44	971.5	26.7	713	S01999	phenylalanine amno
45	540.5	14.8	511	E82228	histidine ammonia-

ALIGNMENTS

RESULT 1

T12749 phenylalanine ammonia-lyase (EC 4.3.1.5) - common sunflower
C:Species: Helianthus annuus (common sunflower)
C:Date: 13 Aug 1999 #sequence_revision 13 Aug 1999 #text_change 29 Jun 2000
C:Accession: T12749
R:Marzegrat, F.A.
Submitted to the EMBL Data Library, February 1998
A:Reference number: 217579
A:Accession: T12749
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-667 <MAX>
A:Cross-references: EMBL:Y12461
C:Gene(s):
A:Gene: PAL
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:197-199/Cross-link: 5-imidazolinone (Ala-Gly) #status predicted
F:198/Modified site: dehydratant (Ser) #status predicted

Query Match	88.0%	Score 3206	DB 2	Length 667
Best Local Similarity	94.4%	Pred. No. 1.1e-197		
Matches 667	Conservative 13	Mismatches 24	Indels 0	Gaps 0
QY	1	MENNHVNCVYVNEI/TKDPL	NGVAAEATGSHIDEVKKWAEERKRVKIGCHITVSQ	60
DB	1	MENGHVNCVYVNEI/TKDPL	NGVAAEATGSHIDEVKKWAEERKRVKIGCHITVSQ	60
QY	61	VATIAAANISIVVYVSHAAWATKASWVMSMNCGISVYTGFTASHPKKG		120
DB	61	VATISAAAGNKKVYVSHAAWATKASWVMSMNCGISVYTGFTASHPKKG		120
QY	121	GALKEELREFLVNGIFGNGTETSHITPHSATKAMIVRINTLQVYSAINFELDEATIKF		180
DB	121	GALQKEELREFLVNGIFGNGTETSHITPHSATKAMIVRINTLQVYSAINFELDEATIKF		180
QY	181	INNNITPCGIPGCTITFASQWVPSYVADLIGKNSAAVGGVIGVIVNAEAPAAAVG		240
DB	181	INNNITPCGIPGCTITFASQWVPSYVADLIGKNSAAVGGVIGVIVNAEAPAAAVG		240
QY	241	GPEFLQPKGGLAVNGTAVGSMASVLPDANVLATSEVISAIFAEVMQCKPPTTHIT		300
DB	241	GPEFLQPKGGLAVNGTAVGSMASVLPDANVLATSEVISAIFAEVMQCKPPTTHIT		300
QY	301	HKRHHIPGJIEAAAIMETITPSSDYVKAQVHEMLPKYKFKUDYALRTSPQDL		360
DB	301	HKLKHHPGJIEAAAIMETITPSSDYVKAQVHEMLPKYKFKUDYALRTSPQDL		360
QY	361	VIRSTKMKEREINSVNQPILIVSNNKALHGGNGGGLPCVSMNHLAIAAIGKIMPA		420
DB	361	VIRSATKMKEREINSVNQPILIVSNNKALHGGNGGGLPCVSMNHLAIAAIGKIMPA		420

```

OY 421 QFSELVDFYNNGLPSNLSCGRNPSLDYGFCKGCEIAMASYCSELOFLANPTVTHNVSADQ 480
DB 421 QFSELVDFYNNGLPSNLSCGRNPSLDYGFCKGCEIAMASYCSELOFLANPTVTHNVSADQ 480
OY 481 HNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLEENNKSTVKNVSOVAK 540
DB 481 HNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLEENNKSTVKNVSOVAK 540
OY 541 KVLTMGVNGELHPSFCECKDLLRVVDREYFAVYIDVCSGTYPLMOKLROVLDHALNNG 600
DB 541 KVLTMGVNGELHPSFCECKDLLRVVDREYFAVYIDVCSGTYPLMOKLROVLDHALNNG 600
OY 601 ETEKNTNPSIFOKIATEFEELKVLIPKFEVGRVAYENDTLSPNRIKACRSYPLRYFV 660
DB 601 ETEKNTNPSIFOKIATEFEELKVLIPKFEVGRVAYENDTLSPNRIKACRSYPLRYFV 660
OY 661 EELG 664
DB 661 EELG 664

```

RESULT 2

S04463
phenylalanine ammonia-lyase (EC 4.3.1.5) - parsley
C:Species: Petroselinum crispum (parsley)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 21-May-1999
C:Accession: S04463
R:Lois, R.; Dietrich, A.; Hahlbrock, K.; Schulz, W.
EMBL J. 8, 1641-1648, 1989
A:Title: A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification
A:Reference number: S04463, MUID:89356633, PMID:2767049
A:Accession: S04463
A:Molecule type: DNA
A:Residues: 1 716 <LOI>
A:Cross-references: EMBL:X15473
C:Genetics:
A:Introns: 136/2
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:203/204/Cross-link: 5-imidazolone (Ala-Gly) #status predicted
F:203/Modified site: denhydroalanine (Ser) #status predicted

Query Match 87.4%; Score 3186.5; DB 2; Length 716;
Best Local Similarity 86.2%; Pred. No. 2.3e-196;
Matches 620; Conservative 46; Mismatches 42; Indels 11; Gaps 4;

```

OY 1 MENCN-----HNGGVNHLCTK--DPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLG 52
DB 1 MENCNATTNCHVNGNMGDMCKTEDEPLWYGTAAEAMTSGHIDEYKKMAEPRKPVYKLG 60
OY 53 GETLTVSOVAGITAAANDSDTYKVELSEARAGYKASSDWYMESNMKGTDSYGVTTGPGAT 112
DB 61 GETLTVSOVAAI-SARDOSGVTVLSEARAGYKASSDWYMDNMNKGTDSYGVTTGPGAT 119
OY 113 SHRRIKOGCALOKELIRPLNAGIFGNCNTHSHLPHSATRAAMIVRINTLLQSYGIRRE 172
DB 120 SHRRIKOGCALOKELIRPLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQSYGIRRE 177
OY 173 ILFAITKFLNNNTIPCLPRLKGTITASGDVPLSYIAGLLTGKRNKAVAGPTGCVLNAFKA 232
DB 178 ILFAITKFLNNNTIPCLPRLKGTITASGDVPLSYIAGLLTGKRNKAVAGPTGCVLNAFKA 237
OY 233 FAAGVGEQFELOPKELALVNGTAVSGMASVLFDAVNLALSEVLISATIAEYMOGK 292
DB 238 FKLAGVEGFELOPKELALVNGTAVSGMASVLFDAVNLALSEVLISATIAEYMOGK 297
OY 293 PEFTDHLTHKILKHPGQIEAAAMEYILDSGDYKAAQKQVHMDPLQKPRQDGYALRTSP 352
DB 298 PEFTDHLTHKILKHPGQIEAAAMEYILDSGDYKAAQKQKHEMDPLQKPRQDGYALRTSP 357
OY 353 QMLGQIEVILKSSITKMLEREINSVNDPLIDVSRNKALHCGNFQGTPIGVSMNTRIALA 412
DB 353 QMLGQIEVILKSSITKMLEREINSVNDPLIDVSRNKALHCGNFQGTPIGVSMNTRIALA 412

```

```

DB 358 QMLGQIEVILKSSITKMLEREINSVNDPLIDVSRNKALHCGNFQGTPIGVSMNTRIALA 417
OY 413 AIGKLMFAQFSELVNDYNNGLPSNLSCGRNPSLDYGFCKGCEIAMASYCSELOFLANPT 472
DB 418 AIGKLMFAQFSELVNDYNNGLPSNLSCGRNPSLDYGFCKGCEIAMASYCSELOFLANPT 477
OY 473 HNVSADQHNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLEENNKSTVKN 532
DB 478 HNVSADQHNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLEENNKSTVKN 537
OY 533 NTVSOVAKKVLITGVNGELHPSFCECKDLLRVVDREYFAVYIDVCSGTYPLMOKLROVL 592
DB 538 NTVSOVAKKVLITGVNGELHPSFCECKDLLRVVDREYFAVYIDVCSGTYPLMOKLROVL 597
OY 593 VDHALLNGETCKNTNPSIFOKIATEFEELKVLIPKFEVGRVAYENDTLSPNRIKACRS 652
DB 598 VDHALLNGETCKNTNPSIFOKIATEFEELKVLIPKFEVGRVAYENDTLSPNRIKACRS 657
OY 653 YPLKFEVGRVAYENDTLSPNRIKACRSYPLRYFV 711
DB 658 YPLKFEVGRVAYENDTLSPNRIKACRSYPLRYFV 716

```

RESULT 3

S48726
phenylalanine ammonia-lyase (EC 4.3.1.5) 3 - parsley
C:Species: Petroselinum crispum (parsley)
C:Date: 18-Oct-1995 #sequence_revision 12-Apr-1996 #text_change 16-Jul-1999
C:Accession: S48726; S56035
R:Appert, C.; Logemann, E.; Hahlbrock, K.; Schmid, J.; Amrhein, N.
Eur. J. Biochem. 225, 491-499, 1994
A:Title: Structural and catalytic properties of the four phenylalanine ammonia-lyase
A:Reference number: S48725; MUID:95010141; PMID:7925471
A:Accession: S48726
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-718 <APP1>
A:Cross-references: EMBL:X81159
R:Appert, C.
submitted to the EMBL Data Library, August 1994
A:Reference number: S56035
A:Accession: S56035
A:Molecule type: mRNA
A:Residues: 1-58, 'E', 60-718 <APP2>
A:Cross-references: EMBL:X81159; NID:9535007, PIDN:CAA57057.1; FID:9535008
C:Genetics:
A:Gene: PAL3
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:204/206/Cross-link: 5-imidazolone (Ala-Gly) #status predicted
F:205/Modified site: denhydroalanine (Ser) #status predicted

Query Match 87.4%; Score 3186.5; DB 2; Length 718;
Best Local Similarity 87.2%; Pred. No. 2.3e-196;
Matches 621; Conservative 46; Mismatches 40; Indels 5; Gaps 4;

```

OY 3 NGNHVGVNDELCTK--DPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLG 60
DB 9 NG-NHNGVDELCTK--DPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLG 67
OY 61 VAGIAAANDSDTYKVELSEARAGYKASSDWYMESNMKGTDSYGVTTGPGATSHRPTKOG 120
DB 68 VAGIAAANDSDTYKVELSEARAGYKASSDWYMDNMNKGTDSYGVTTGPGATSHRPTKOG 126
OY 121 GALKKELIRFLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQSYGIRREILFAITK 179
DB 127 GALKKELIRFLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQSYGIRREILFAITK 186
OY 180 FLNNNTIPCLPRLKGTITASGDVPLSYIAGLLTGKRNKAVAGPTGCVLNAEAFAAAGVE 239
DB 187 FLNNNTIPCLPRLKGTITASGDVPLSYIAGLLTGKRNKAVAGPTGCVLNAEAFAAAGVE 246
OY 240 GPFELQPKKELALVNGTAVSGMASVLFDAVNLALSEVLISATIAEYMOGKPEFTDHL 299

```

```

Db 247 GCFEELQPKKGLALVNGTAVGSGMASMVLFRANILAVAFVMSALFAVYMQCKPFFTHI 306
Qy 300 TIRIKTHIFVQTEADAMEYILDSSYVKAAGVHEMLPLQKRCCEPVATLSFQALSPJ 354
Db 307 THRIKTHIFQTEADAMEYILDSSYVKAAGVHEMLPLQKRCCEPVATLSFQALSPJ 366
Qy 360 EVIRSSTKMEREINSVNPNPLIDVSRNKAIHGMNDPLQKRCCEPVATLSFQALSPJ 419
Db 367 EVIRSSTKMEREINSVNPNPLIDVSRNKAIHGMNDPLQKRCCEPVATLSFQALSPJ 426
Qy 420 AQSSELVNPNPNPLIDVSRNKAIHGMNDPLQKRCCEPVATLSFQALSPJ 479
Db 427 AQSSELVNPNPNPLIDVSRNKAIHGMNDPLQKRCCEPVATLSFQALSPJ 486
Qy 480 QINDVNSLALISAKTAEVNDILKLMSSITVALVQSIDLKHLEBNKSYVKNVSVYA 539
Db 487 QINDVNSLALISAKTAEVNDILKLMSSITVALVQSIDLKHLEBNKSYVKNVSVYA 546
Qy 540 KKVITMVGCELHPSHPCCKDLKRVDRKVFAYIDVSCGTYPLMOKLQVLDVHALNN 599
Db 547 KKVITMVGCELHPSHPCCKDLKRVDRKVFAYIDVSCGTYPLMOKLQVLDVHALNN 606
Qy 600 GETHKNTNISIFOKIAFEEELKVLIPKVEGVRIAYENDTISINRIKACRSYPLRYV 659
Db 607 GETHKNTNISIFOKIAFEEELKVLIPKVEGVRIAYENDTISINRIKACRSYPLRYV 666
Qy 660 REELGRGFTGKVTSPGEEFDPVFTAMKQGIIDPLELCTGTCVNMPTPTG 711
Db 667 REELGRGFTGKVTSPGEEFDPVFTAMKQGIIDPLELCTGTCVNMPTPTG 718

```

RESULT 4

S48725
phenylalanine ammonia-lyase (EC 4.3.1.5) 2 - parsley
C:Species: Petroselinum crispum (parsley)
C>Date: 28 Oct 1995 #sequence_revision 10 Nov 1995 #excl_change 1b-Jul-1999
C:Accession: S48725
R:Appert, C.; Lohmann, E.; Hahnbach, K.; Schulz, J.; Amrhein, N.
Eur. J. Biochem. 225, 491-499, 1994
A:Title: Structural and catalytic properties of the four phenylalanine ammonia-lyase iso-
A:Reference number: S48725, MIMD:95010141, PMID:7925471
A:Accession: S48725
A:Status: nucleic acid sequence not shown
A:Residues: 1-716 - APP.
A:Molecule type: mRNA
A:Cross-references: EMBL:X81158; NID:q534892; PIRN:CMA57056.1; PIR:q534893
C:Genetics:
A:Gene: PAL2
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:202-204/cross-link: 5 imidazolinone (Ala-Gly) #status predicted
F:203/Modified site: dehydroalanine (Ser) #status predicted

Query Match 87.4%, Score 3185.5, DB 2; Length 716;
Best Local Similarity 86.2%, Pred. No. 26e-196;
Matches 620; Conservative 45; Mismatches 43; Indels 11; Gaps 4;

```

Qy 1 MONGA-----HNGVYNEICIK--DPLNMGVAAEALTGSHLDEYKKVAEERKPVKLG 52
Db 1 MONGAITNGCHVNGCMQDPCKTIDPLVWGIAAAMTGSHTLDEYKKVAEERKPVKLG 60
Qy 53 GFTLTGQVAGTAANDSDTYKVELSEAPAVKASSDMWESMKNKGTDSYVTTGFGAT 112
Db 61 GFTLTGQVAGTAANDSDTYKVELSEAPAVKASSDMWESMKNKGTDSYVTTGFGAT 119
Qy 113 SHRTKGGALQKELIRFLNAGIFGCGTETSHLPHSATRAAMVRIINTLQGYSGIRPE 172
Db 120 SHRTKGGALQKELIRFLNAGIFGCGTETSHLPHSATRAAMVRIINTLQGYSGIRPE 177
Qy 173 ILEATITLNNNTTLCPLRGITTSAGDLVPLSYTAGLLTGPPNKAAGVPTGEVLAERKA 232
Db 178 ILEATITLNNNTTLCPLRGITTSAGDLVPLSYTAGLLTGPPNKAAGVPTGEVLAERKA 237

```

```

Qy 233 FAAGVNDSTHETLQKRCCEPVATLSFQALSPJ 292
Db 238 FAAGVNDSTHETLQKRCCEPVATLSFQALSPJ 297
Qy 293 PEPIDHITKRIKTHIFQTEADAMEYILDSSYVKAAGVHEMLPLQKRCCEPVATLSFQALSPJ 352
Db 298 PEPIDHITKRIKTHIFQTEADAMEYILDSSYVKAAGVHEMLPLQKRCCEPVATLSFQALSPJ 357
Qy 353 QINDVNSLALISAKTAEVNDILKLMSSITVALVQSIDLKHLEBNKSYVKNVSVYA 412
Db 358 QINDVNSLALISAKTAEVNDILKLMSSITVALVQSIDLKHLEBNKSYVKNVSVYA 417
Qy 413 AQSSELVNPNPNPLIDVSRNKAIHGMNDPLQKRCCEPVATLSFQALSPJ 472
Db 418 AQSSELVNPNPNPLIDVSRNKAIHGMNDPLQKRCCEPVATLSFQALSPJ 477
Qy 473 NHOVSARHODVNSLALISAKTAEVNDILKLMSSITVALVQSIDLKHLEBNKSYVKNVSVYA 532
Db 478 NHOVSARHODVNSLALISAKTAEVNDILKLMSSITVALVQSIDLKHLEBNKSYVKNVSVYA 537
Qy 533 NTVSVAKKVITMVGCELHPSHPCCKDLKRVDRKVFAYIDVSCGTYPLMOKLQVLDVHALNN 592
Db 538 NTVSVAKKVITMVGCELHPSHPCCKDLKRVDRKVFAYIDVSCGTYPLMOKLQVLDVHALNN 597
Qy 593 VDHALNNGEFTGKVTSPGEEFDPVFTAMKQGIIDPLELCTGTCVNMPTPTG 652
Db 598 VDHALNNGEFTGKVTSPGEEFDPVFTAMKQGIIDPLELCTGTCVNMPTPTG 657
Qy 653 YPLRYVREELQPKKGLALVNGTAVGSGMASMVLFRANILAVAFVMSALFAVYMQCKPFFTHI 711
Db 658 YPLRYVREELQPKKGLALVNGTAVGSGMASMVLFRANILAVAFVMSALFAVYMQCKPFFTHI 716

```

RESULT 5

T14295
phenylalanine ammonia-lyase (EC 4.3.1.5) - carrot
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #excl_change 21-Jul-2000
C:Accession: T14295
R:Takeuchi, Y.; Otsuki, Y.; Yoshida, K.
Photochem. Photobiol. 66, 464-470, 1997
A:Title: Action spectrum for induction of promoter activity of phenylalanine ammonia-
A:Reference number: 217966, MIMD:97478956, PMID:9337617
A:Accession: T14295
A:Status: Preliminary, translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-708 - TRAK>
A:Cross-references: EMBL:D85850; NID:q2618589; PIRN:BA42367.1; PIR:q2618590
C:Experimental source: strain Kuradagousun
C:Genetics:
A:Gene: PAL1
A:Introns: 126/2
C:Function:
A:Description: catalyzes the deamination of L-phenylalanine to trans-cinnamic acid
A:Pathway: phenylpropanoid metabolism
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:14-196/cross-link: 5 imidazolinone (Ala-Gly) #status predicted
F:195/Modified site: dehydroalanine (Ser) #status predicted

Query Match 87.0%, Score 3171.5, DB 2; Length 708;
Best Local Similarity 87.2%, Pred. No. 2e-195;
Matches 619; Conservative 41; Mismatches 43; Indels 7; Gaps 3;

```

Qy 2 ENCHRVNGVYNEICIK--DPLNMGVAAEALTGSHLDEYKKVAEERKPVKLG 60
Db 4 ENKRVYVIG--NGICMOKDPLNMGMAAKALTGSHLDEYKKVAEERKPVKLG 61
Qy 61 VAGTAANDSDTYKVELSEAPAVKASSDMWESMKNKGTDSYVTTGFGATSHRTKGG 120
Db 62 VAGTAANDSDTYKVELSEAPAVKASSDMWESMKNKGTDSYVTTGFGATSHRTKGG 117

```

```

QY 121 GATGKELIRFLNAGIFNGTETSHLPASATRAAMIVKINTLLDGGSGIRFEITLAIKFE 180
Db 118 CAIQKELIRFLNAGIFNGTETSHLPASATRAAMIVKINTLLDGGSGIRFEITLAIKFE 177
QY 181 LNNITPCLPLRGITITASGDLVPLSYTAGLLTGRPNKAVGPTGEVIMAKKAPAAAGVGC 240
Db 178 LNNITPCLPLRGITITASGDLVPLSYTAGLLTGRPNKAVGPTGEVIMAKKAPAAAGVGC 237
QY 241 GFEELQKREGIALVNGTAVSGMASVYLFDANVALLSEVLASIPANVMOGKEPFDHLT 300
Db 238 GFEELQKREGIALVNGTAVSGMASVYLFDANVALLSEVLASIPANVMOGKEPFDHLT 297
QY 301 HKLHHHPQIEFAAIMEYILDGSDYKAAOKVHEMDPLQKPKODRYALRTSPQMLGPOIE 360
Db 298 HKLHHHPQIEFAAIMEYILDGSDYKAAOKVHEMDPLQKPKODRYALRTSPQMLGPOIE 357
QY 361 VIRSTKMIEREINSVNDNPLIDVSRNKAHLHGNGFOGTPIGVSMONTPLAIAIGKLMA 420
Db 358 VIRSTKMIEREINSVNDNPLIDVSRNKAHLHGNGFOGTPIGVSMONTPLAIAIGKLMA 417
QY 421 QFSELVNDYFVNGGLPSNLSGGRNPGLDYGFKGGEIAMASYCELOFLANPVTNHVQSAQ 480
Db 418 QFSELVNDYFVNGGLPSNLSGGRNPGLDYGFKGGEIAMASYCELOFLANPVTNHVQSAQ 477
QY 481 HNUVNSLGLISARKTAEAVDILKLMSSYVALCOSIDLRLLENNKSTVKNTVSQVAK 540
Db 478 HNUVNSLGLISARKTAEAVDILKLMSSYVALCOSIDLRLLENNKSTVKNTVSQVAK 537
QY 541 KVLTMGVNGLHPSPFCEKDLRVYDREYFAYIDVCGSTYPLMOKLNOVLDHALNNG 600
Db 538 KVLTMGVNGLHPSPFCEKDLRVYDREYFAYIDVCGSTYPLMOKLNOVLDHALNNG 597
QY 601 ETEKNTNTSIFOKIATFEELKVLPLKEVEGVRIAYENDTISIPNRIKACRSYPLRFAV 660
Db 598 ETEKNTNTSIFOKIATFEELKVLPLKEVEGVRIAYENDTISIPNRIKACRSYPLRFAV 657
QY 661 EELRGFTLGKVTSPGEEFDRVFTAMCKGQIIDLPLECLGGNNGEPLPI 710
Db 658 EELRGFTLGKVTSPGEEFDRVFTAMCKGQIIDLPLECLGGNNGEPLPI 707

RESULT 6
S66343
phenylalanine ammonia-lyase (EC 4.3.1.5) 1 - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 16-Jul-1999
C:Accession: S66343
R:Fukasawa-Akada, T.; Kung, S.; Watson, J.C.
Plant Mol Biol. 30, 711-722, 1996
A:Title: Phenylalanine ammonia-lyase gene structure, expression, and evolution in Nicotiana glauca
A:Reference number: S66343; MUID:96194455; PMID:8624404
A:Accession: S66343
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-715 <FUK>
A:Cross-references: EMBL:M84466; NID:g170349; PION:AAA34122.1; PTD:g170350
C:Genetics:
A:Gene: PAL1
A:Introns: 13/2
A:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:201-203/cross-link: 5-imidazolone (Ala-Gly) #status predicted
F:202/modified site: dehydroalanine (Ser) #status predicted

Query Match 86.4%; Score 3149; DB 2; Length 715;
Best Local Similarity 85.3%; Pred. No. 5,7e-194;
Matches 610; Conservative 45; Mismatches 56; Indels 4; Gaps 3;

QY 1 MEMONHNGVNV-ELGKIK--DPLNMGVAALALGSHLDYKKKVAARPKRPVYKIGGTHIR 57
Db 1 MASGHNHNGGNEFLCKKSDAPLNMWMAARSLGSHLDYKKKVAARPKRPVYKIGGTHIR 60
QY 58 VSQVAGIAAANDS-DIVKVELSEARAGVAKASDWMESMKNKGTDSYGVTTGGATSHRR 116

```

```

Db 61 VAQVAAIAAVNDKSNKGVKVELSEARAGVAKASDWMESMKNKGTDSYGVTTGGATSHRR 120
QY 117 TKGGALQKELIRFLNAGIFNGTETSHLPASATRAAMIVKINTLLDGGSGIRFEITLAIKFE 176
Db 121 TKGGALQKELIRFLNAGIFNGTETSHLPASATRAAMIVKINTLLDGGSGIRFEITLAIKFE 180
QY 177 ITRFLNNITPCLPLRGITITASGDLVPLSYTAGLLTGRPNKAVGPTGEVIMAKKAPAAAGVGC 240
Db 181 ITRFLNNITPCLPLRGITITASGDLVPLSYTAGLLTGRPNKAVGPTGEVIMAKKAPAAAGVGC 237
QY 237 GVEGFEELQKREGIALVNGTAVSGMASVYLFDANVALLSEVLASIPANVMOGKEPFDHLT 300
Db 241 GVEGFEELQKREGIALVNGTAVSGMASVYLFDANVALLSEVLASIPANVMOGKEPFDHLT 297
QY 297 DHLTTHKHHHPQIEFAAIMEYILDGSDYKAAOKVHEMDPLQKPKODRYALRTSPQMLGPOIE 360
Db 301 DHLTTHKHHHPQIEFAAIMEYILDGSDYKAAOKVHEMDPLQKPKODRYALRTSPQMLGPOIE 357
QY 357 POIEVIRSTKMIEREINSVNDNPLIDVSRNKAHLHGNGFOGTPIGVSMONTPLAIAIGKLMA 420
Db 361 POIEVIRSTKMIEREINSVNDNPLIDVSRNKAHLHGNGFOGTPIGVSMONTPLAIAIGKLMA 417
QY 417 LMFQFSELVNDYFVNGGLPSNLSGGRNPGLDYGFKGGEIAMASYCELOFLANPVTNHVQ 480
Db 421 LMFQFSELVNDYFVNGGLPSNLSGGRNPGLDYGFKGGEIAMASYCELOFLANPVTNHVQ 478
QY 477 SABOHNDVNSLGLISARKTAEAVDILKLMSSYVALCOSIDLRLLENNKSTVKNTVSQVAK 540
Db 481 SABOHNDVNSLGLISARKTAEAVDILKLMSSYVALCOSIDLRLLENNKSTVKNTVSQVAK 537
QY 537 QVARKVLTMGVNEFLHPSPFCEKDLRVYDREYFAYIDVCGSTYPLMOKLNOVLDHALNNG 600
Db 541 QVARKVLTMGVNEFLHPSPFCEKDLRVYDREYFAYIDVCGSTYPLMOKLNOVLDHALNNG 597
QY 597 LNNETKNTNTSIFOKIATFEELKVLPLKEVEGVRIAYENDTISIPNRIKACRSYPLRFAV 660
Db 601 LNNETKNTNTSIFOKIATFEELKVLPLKEVEGVRIAYENDTISIPNRIKACRSYPLRFAV 657
QY 657 RFVREELGKGFPLGKVTSPGEEFDRVFTAMCKGQIIDLPLECLGGNNGEPLPI 711
Db 661 RFVREELGKGFPLGKVTSPGEEFDRVFTAMCKGQIIDLPLECLGGNNGEPLPI 707

RESULT 7
T03663
phenylalanine ammonia-lyase (EC 4.3.1.5) - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Jul-1999
C:Accession: T03663
R:Belleguind, L.; Kohfritsch, O.; Fritig, B.; Legrand, M.
Plant Physiol. 106, 877-886, 1994
A:Title: Phenylalanine ammonia-lyase in tobacco. Molecular cloning and gene expression
A:Reference number: T03663; MUID:95125127; PMID:7824656
A:Accession: T03663
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-712 <PPL>
A:Cross-references: EMBL:X78263; NID:g633596; PION:CAA55075.1; PTD:g633597
A:Experimental source: cultivar Samsun NN; tissue-type leaves after infection by tobacco etch virus
A:Gene: PAL
C:Function:
A:Description: catalyzes the deamination of L-phenylalanine to trans-cinnamic acid
A:Note: strongly induced during the hypersensitive reaction to TMV or to a fungal pathogen
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:198-200/cross-link: 5-imidazolone (Ala-Gly) #status predicted
F:199/modified site: dehydroalanine (Ser) #status predicted

Query Match 85.6%; Score 3118.5; DB 2; Length 712;
Best Local Similarity 86.0%; Pred. No. 5,2e-192;
Matches 602; Conservative 38; Mismatches 59; Indels 1; Gaps 1;

```

QY 13 ELCTK-DPLMNGVAEALTGSHLDEYKKWAEFRKRPVKLTGSELTIVSQVAGIAAANDSD 71
 Db 13 DPCVKVPLTMMHMAOSI,KCSHLDYKKWAEFRKRPVKLTGSELTIVSQVAGIAAANDSD 72
 QY 72 TVKVELSEARAVYKASSEWMEKMGTDGSGVTTGTFATSHPTKPGALOKELIRL 131
 Db 73 TVKVELSEARAVYKASSEWMEKMGTDGSGVTTGTFATSHPTKPGALOKELIRL 132
 QY 132 NAGTFCGCTSHLPHSATKAAMIVKINTLLOGYSGLREHLEALITKLLNNITPCLPL 191
 Db 133 NAGTFCGCTSHLPHSATKAAMIVKINTLLOGYSGLREHLEALITKLLNNITPCLPL 192
 QY 192 RGTITAGSDIPLSTYAGLTGTPRNSKAVPTGTEVLNAEKAFAAIVEGFELOPKRCL 251
 Db 193 RGTITAGSDIPLSTYAGLTGTPRNSKAVPTGTEVLNAEKAFAAIVEGFELOPKRCL 252
 QY 252 ALVNGTAVSGMAAMVLEFANVLALESEVLSAIFAEVMOGRPEFTDHLTKLKHHPQTE 311
 Db 253 ALVNGTAVSGMAAMVLEFANVLALESEVLSAIFAEVMOGRPEFTDHLTKLKHHPQTE 312
 QY 312 AAAIMEYLLDGSIVYKAAQVHMDPLQKPKQKVAALRTSPQWLGQVIVKSSITKIMHR 371
 Db 313 AAAIMEYLLDGSIVYKAAQVHMDPLQKPKQKVAALRTSPQWLGQVIVKSSITKIMHR 372
 QY 372 EINSVNDNPLIDVSRNKALHG:NFG:TP:IGVSMNTPLAIAALGKMFQSELYVDFYN 431
 Db 373 EINSVNDNPLIDVSRNKALHG:NFG:TP:IGVSMNTPLAIAALGKMFQSELYVDFYN 432
 QY 432 NGDPSNLISGGRNPSLDYGEKGEIAMAASYSCELOFLANPVTNHVSAEDHNDVSLGL 491
 Db 433 NGDPSNLISGGRNPSLDYGEKGEIAMAASYSCELOFLANPVTNHVSAEDHNDVSLGL 492
 QY 492 SARKTAFAVDILKIMSTYIVAIKOSIDLRLHFNKMTYKNTVNSQVAKKVIIMVGNCL 551
 Db 493 SARKTAFAVDILKIMSTYIVAIKOSIDLRLHFNKMTYKNTVNSQVAKKVIIMVGNCL 552
 QY 552 HPSRCEKDLRYVDREYFAYIDVCSGTYPLMOKLROYLDHALNGETKNTSIF 611
 Db 553 HPSRCEKDLRYVDREYFAYIDVCSGTYPLMOKLROYLDHALNGETKNTSIF 612
 QY 612 OKITFEELKVLIPKEVEGRIVAYENDTSLIPNRKAKRSPRYAFVEBELCPGFLTGE 671
 Db 613 OKITFEELKVLIPKEVEGRIVAYENDTSLIPNRKAKRSPRYAFVEBELCPGFLTGE 672
 QY 672 KVTSPGEFDRVFTAMCKQOIIDPLHCLQCMNGEPPLPIC 711
 Db 673 KVTSPGEFDRVFTAMCKQOIIDPLHCLQCMNGEPPLPIC 712

RESULT 8

J02265
 C:Species: *Populus trichocarpa* x *Populus deltoides* (western balsam poplar x cottonwood)
 C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 28-May-1999
 C:Accession: J02265
 R:Subramaniam, R.; Reinold, S.; Mollitor, E.K.; Douglas, C.J.
 A:Title: Structure, inheritance, and expression of hybrid poplar (*Populus trichocarpa* x *Populus deltoides*)
 A:Reference number: J02265, PMID:94151434, PMID:8108506
 A:Accession: J02265
 A:Molecule type: mRNA
 A:Residues: 1-715 <SRV>
 A:Cross references: GH:11747; NID:9169453; PUDN:AAA3805.1; PUD:9169454
 A:Note: The authors translated the codon AAG for residue 331 as Glu and AAG for residue 332 as Lys. The authors also translated the codon AAG for residue 331 as Glu and AAG for residue 332 as Lys.
 C:Comment: The enzyme plays a role in linking primary metabolism to phenylpropanoid metabolism.
 C:Keywords: ammonia-lyase, carbon-nitrogen lyase
 F:201-203/Cross-link: 5 imidazolinone (Ala Gly) #status predicted
 F:202/Modified site: dehydroalanine (Ser) #status predicted

Query Match 85.5%; Score 3116, FR 2; Length 715;
 Best Local Similarity 84.1%; Pred. No. 7, 5e-192;

Matches 599; Conservative 56; Mismatches 53; Indels 4; Gaps 3;
 QY 2 ENGHVNGVYNEHCT--KDFLNMGVAAEALTGSHLDEYKKWAEFRKRPVKLTGSELTIVSQ 59
 Db 6 KNG-YONGSLFSLICVNDPDLTSMGVAAEALTGSHLDEYKKWAEFRKRPVKLTGSELTIVSQ 64
 QY 60 UVAIAAANSLDVKVELSEARAVYKASSEWMEKMGTDGSGVTTGTFATSHPTKPGALOKELIRL 119
 Db 65 QVAST-AGHTGVYKVSLSAPPGYKASSDWMKMGTDGSGVTTGTFATSHPTKPGALOKELIRL 123
 QY 120 GVALOKELIRLPLNAGTFCGCTSHLPHSATKAAMIVKINTLLOGYSGLREHLEALITKLLNNITPCLPL 179
 Db 124 GVALOKELIRLPLNAGTFCGCTSHLPHSATKAAMIVKINTLLOGYSGLREHLEALITKLLNNITPCLPL 183
 QY 180 PLANNITPCLPLNAGTFCGCTSHLPHSATKAAMIVKINTLLOGYSGLREHLEALITKLLNNITPCLPL 239
 Db 184 PLANNITPCLPLNAGTFCGCTSHLPHSATKAAMIVKINTLLOGYSGLREHLEALITKLLNNITPCLPL 243
 QY 240 GGFELDPKRYALVNGTAVSGMAAMVLEFANVLALESEVLSAIFAEVMOGRPEFTDHLTKLKHHPQTE 299
 Db 244 GGFELDPKRYALVNGTAVSGMAAMVLEFANVLALESEVLSAIFAEVMOGRPEFTDHLTKLKHHPQTE 303
 QY 300 THKLKHHPQTEAAAIMEYLLDGSIVYKAAQVHMDPLQKPKQKVAALRTSPQWLGQVIVKSSITKIMHR 359
 Db 304 THKLKHHPQTEAAAIMEYLLDGSIVYKAAQVHMDPLQKPKQKVAALRTSPQWLGQVIVKSSITKIMHR 363
 QY 360 EVIKSSIKMTPEKINSVNDNPLIDVSRNKALHG:NFG:TP:IGVSMNTPLAIAALGKMFQSELYVDFYN 419
 Db 364 EVIKSSIKMTPEKINSVNDNPLIDVSRNKALHG:NFG:TP:IGVSMNTPLAIAALGKMFQSELYVDFYN 423
 QY 420 AGFSELVNDNPLIDVSRNKALHG:NFG:TP:IGVSMNTPLAIAALGKMFQSELYVDFYN 479
 Db 424 AGFSELVNDNPLIDVSRNKALHG:NFG:TP:IGVSMNTPLAIAALGKMFQSELYVDFYN 483
 QY 480 OHNDVNSLGLSARKTAFAVDILKIMSTYIVAIKOSIDLRLHFNKMTYKNTVNSQVAKKVIIMVGNCL 539
 Db 484 OHNDVNSLGLSARKTAFAVDILKIMSTYIVAIKOSIDLRLHFNKMTYKNTVNSQVAKKVIIMVGNCL 543
 QY 540 KVTSPGEFDRVFTAMCKQOIIDPLHCLQCMNGEPPLPIC 599
 Db 544 KVTSPGEFDRVFTAMCKQOIIDPLHCLQCMNGEPPLPIC 603
 QY 600 GETEKNTSIFOKITFEELKVLIPKEVEGRIVAYENDTSLIPNRKAKRSPRYAFVEBELCPGFLTGE 659
 Db 604 GETEKNTSIFOKITFEELKVLIPKEVEGRIVAYENDTSLIPNRKAKRSPRYAFVEBELCPGFLTGE 663
 QY 660 REHLEALITKLLNNITPCLPLNAGTFCGCTSHLPHSATKAAMIVKINTLLOGYSGLREHLEALITKLLNNITPCLPL 711
 Db 664 REHLEALITKLLNNITPCLPLNAGTFCGCTSHLPHSATKAAMIVKINTLLOGYSGLREHLEALITKLLNNITPCLPL 715

RESULT 9

J01858
 C:Species: *Nicotiana glauca* (common tobacco)
 C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 20-Jun-2000
 C:Accession: J01858
 R:Taguchi, G.; Sharan, M.; Gunda, K.; Yanojima, K.; Shimoda, M.; Hayashida, N.; O
 J. Plant biochem. and biotech. 7, 79-84, 1998
 A:Title: Effect of methyl jasmonate and elicitor on PAL gene expression in tobacco callus
 A:Reference number: Z14446
 A:Accession: J01858
 A:Status: preliminary, translated from GH/EMBL/JDBJ
 A:Molecule type: DNA
 A:Residues: 1-712 <TAG>
 A:Cross references: EMBL:AB06199; PUDN:BA22447.1
 A:Experimental source: strain Bright yellow; tissue type callus
 C:Genetics:
 A:Gene: palA
 A:Introns: 130/2
 C:Function:
 A:Description: catalyzes the decarboxylation of L-phenylalanine to trans-cinnamic acid
 C:Superfamily: histidine ammonia-lyase

C:Keywords: ammonia-lyase; carbon-nitrogen lyase
 F:198-200/Cross-link: 5-Imidazolinone (Ala-Gly) #status predicted
 F:199/Modified site: dehydroalanine (Ser) #status predicted

Query Match 85.3%; Score 3109.5; DB 2; Length 712;
 Best Local Similarity 86.1%; Pred. No. 1.9e-191;
 Matches 603; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 13 ELCTK-DPLNMGVAAEALTGSHLDEVKKMAVEPRKPVYKLGGLTLYSOVAATAANDSD 71
 DB 13 DPCMKVDPNMMKADSLKSHLDEVKKMAVEPRKPVYKLGGLTLYSOVAATAANDSD 72
 QY 72 TVKVELSEARAGVASSDWMVSNKKTDSYVTTGATSHRRTKOGALOKELRPL 131
 DB 73 TVKVELSEARAGVASSDWMVSNKKTDSYVTTGATSHRRTKOGALOKELRPL 132
 QY 132 NGITGNGTSHITPHSATRAMIVRINTLLQGSIGRITELATKFLNNITPCLPL 191
 DB 133 NGVGVNGTSHITPHSATRAMIVRINTLLQGSIGRITELATKFLNNITPCLPL 192
 QY 192 RGTITASDPLVPLSYIAGLTGKPNKAVGPTGEVLNAEKFAAGVGGFFELQPEGL 251
 DB 193 RGTITASDPLVPLSYIAGLTGKPNKAVGPTGEVLNAEKFAAGVGGFFELQPEGL 252
 QY 252 ALVNGTAVGSGMASVLPDANVIALISEVLSAIFAEVWQCKPEFTDHLTKLKHHPQIE 311
 DB 253 ALVNGTAVGSGMASVLPDANVIALISEVLSAIFAEVWQCKPEFTDHLTKLKHHPQIE 312
 QY 312 AAIVIEYITDSDYKAAOKVHEMDPLQPKODRYALRTSPWQGTQIEVRSSTKMER 371
 DB 313 AAIVIEYITDSDYKAAOKVHEMDPLQPKODRYALRTSPWQGTQIEVRSSTKMER 372
 QY 372 EINSVNDNPLIDVSRNKKALHGGNFGPTIGVSMONTRILATIAIGKLMFAQFSELVNDPYN 431
 DB 373 EINSVNDNPLIDVSRNKKALHGGNFGPTIGVSMONTRILATIAIGKLMFAQFSELVNDPYN 432
 QY 432 NGLPBNLSGGRPSIDYGFPGKGTAMASGSELOFLANPVYNHVSQAQHNQVNSIGLI 491
 DB 433 NGLPBNLSGGRPSIDYGFPGKGTAMASGSELOFLANPVYNHVSQAQHNQVNSIGLI 492
 QY 492 SARKEAEVNDILKLMSSYIYVALCQSIDRLHLEENKSTVANTVSQVAKVLTGAVGEL 551
 DB 493 SARKEAEVNDILKLMSSYIYVALCQSIDRLHLEENKSTVANTVSQVAKVLTGAVGEL 552
 QY 552 HPSRCEKDLAVNDREYFAVIDVCSGTPLMOKLEQVIVDHALNNGETKNTSTIF 611
 DB 553 HPSRCEKDLAVNDREYFAVIDVCSGTPLMOKLEQVIVDHALNNGETKNTSTIF 612
 QY 612 OKIATFEELKAVLIPKVEGVRAVENDTLSIPNRIKACRSPLYRIVREELGKPLTGE 671
 DB 613 OKIATFEELKAVLIPKVEGVRAVENDTLSIPNRIKACRSPLYRIVREELGKPLTGE 672
 QY 672 KVTSPGEEDRVFTAMCKGQIIDPLECLGSGNNGEPLPIC 711
 DB 673 KVTSPGEEDRVFTAMCKGQIIDPLECLGSGNNGEPLPIC 712

RESULT 10

JC5872
 phenylalanine ammonia-lyase (EC 4.3.1.5) 1 - Lithospermum erythrorhizon
 C:Species: Lithospermum erythrorhizon
 C:Date: 09-Mar-1998 #sequence_revision 09-Mar-1998 #text_change 20-Jun-2000
 C:Accession: JC5872; PC4506
 R:Yazaki, K.; Kataoka, M.; Honda, G.; Severin, K.; Heide, L.
 Biosci. Biotechnol. Biochem. 61, 1995-2003, 1997
 A:Title: cDNA cloning and gene expression of phenylalanine ammonia-lyase in Lithospermum
 A:Reference number: JC5872, MUID 94101947, PMID 9418988
 A:Accession: JC5872
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-710 <YALI>
 A:Cross-references: DBJ:DB3075, NID g2911123; P10N PAA24928 1; P10N g2911122
 A:Accession: PC4506

A:Molecule type: protein
 A:Residues: 348-479 -YAA2-
 C:Comment: This enzyme catalyzes the first committed step in multibranched phenylprop
 s of roots to protect them against pathogens in the soil
 C:Superfamily: histidine ammonia-lyase
 C:Keywords: ammonia-lyase; carbon-nitrogen lyase
 F:196-198/Cross-link: 5-Imidazolinone (Ala-Gly) #status predicted
 F:197/Modified site: dehydroalanine (Ser) #status predicted

Query Match 84.5%; Score 3077.5; DB 2; Length 710;
 Best Local Similarity 85.5%; Pred. No. 2.2e-189;
 Matches 594; Conservative 53; Mismatches 59; Indels 5; Gaps 2;

QY 1 MENCNHHVQVNEICIKPLNMVAAEALTGSHLDEVKKMAVEPRKPVYKLGGLTLYVSG 60
 DB 5 VENG---NGKIMEPCMKPLNMEVASESKSHLDEVKKMAVEPRKPVYLAACKTLIDQ 61
 QY 61 VAGTAAANSTVYKVEI SEAPACVAKASQWVMSNKKSTVSVCTTGPATSHRRKQK 120
 DB 62 VAAIARDQ--VTVELAEAREGVKASQWVMSNKKSTVSVCTTGPATSHRRKQK 119
 QY 121 GALKKELKFLNAGIFKNGTSTHTPHSATRAMIVRINTLLQGSIGRITELATKFL 180
 DB 120 GALKKELKFLNAGIFKNGTSTHTPHSATRAMIVRINTLLQGSIGRITELATKFL 179
 QY 181 INNNTPLPLRGITITASGDLVPLSYIAGLTGKPNKAVGPTGEVLNAEKFAAGVGGFFELQPEGL 240
 DB 180 INNNTPLPLRGITITASGDLVPLSYIAGLTGKPNKAVGPTGEVLNAEKFAAGVGGFFELQPEGL 239
 QY 241 GFPELOPEKGLALVNGTAVGSGMASVLPDANVIALISEVLSAIFAEVWQCKPEFTDHLT 300
 DB 240 GFPELOPEKGLALVNGTAVGSGMASVLPDANVIALISEVLSAIFAEVWQCKPEFTDHLT 299
 QY 301 HKLKHHPQIEFAAIMEYITDSDYKAAOKVHEMDPLQPKODRYALRTSPWQGTQIEVRSSTKMER 360
 DB 300 HKLKHHPQIEFAAIMEYITDSDYKAAOKVHEMDPLQPKODRYALRTSPWQGTQIEVRSSTKMER 359
 QY 361 VIRSTKMIEREINSVNDNPLIDVSRNKKALHGGNFGPTIGVSMONTRILATIAIGKLMFA 420
 DB 360 VIRSTKMIEREINSVNDNPLIDVSRNKKALHGGNFGPTIGVSMONTRILATIAIGKLMFA 419
 QY 421 QHSELVNPYNNQGPSNLSGGKNSLDYGFPGKGTAMASGSELOFLANPVYNHVSQAQHNQVNSIGLI 480
 DB 420 QHSELVNPYNNQGPSNLSGGKNSLDYGFPGKGTAMASGSELOFLANPVYNHVSQAQHNQVNSIGLI 479
 QY 481 HNGDVNSIGLISARKTAEVNDILKLMSSYIYVALCQSIDRLHLEENKSTVANTVSQVAKVLTGAVGEL 540
 DB 480 HNGDVNSIGLISARKTAEVNDILKLMSSYIYVALCQSIDRLHLEENKSTVANTVSQVAKVLTGAVGEL 539
 QY 541 KVLTVGVNGELHPSRCEKDLAVNDREYFAVIDVCSGTPLMOKLEQVIVDHALNNGEPLPIC 600
 DB 540 KVLTVGVNGELHPSRCEKDLAVNDREYFAVIDVCSGTPLMOKLEQVIVDHALNNGEPLPIC 599
 QY 601 ETEKNTNISIPKTIATFEELKAVLIPKVEGVRAVENDTLSIPNRIKACRSPLYRIVREELGKPLTGE 660
 DB 600 ETEKNTNISIPKTIATFEELKAVLIPKVEGVRAVENDTLSIPNRIKACRSPLYRIVREELGKPLTGE 659
 QY 661 EELGKGLTGERKVTSPGEEEDRVFTAMCKGQIIDPLECLGSGNNGEPLPIC 711
 DB 660 EELGKGLTGERKVTSPGEEEDRVFTAMCKGQIIDPLECLGSGNNGEPLPIC 710

RESULT 11

S25303
 phenylalanine ammonia-lyase (EC 4.3.1.5) - garden pea
 C:Species: Pisum sativum (garden pea)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb 1994 #text_change 20-Jun-2000
 C:Accession: S25303
 R:Kawamura, S.; Yamada, T.; Tanaka, Y.; Sriptraeritsak, P.; Kato, H.; Ichinose, Y.; Ka
 Plant Mol. Biol. 20, 167-170, 1992
 A:Title: Molecular cloning of phenylalanine ammonia-lyase cDNA from Pisum sativum.
 A:Reference number: S25303, MUID:92385763, PMID:1515609
 A:Accession: S25303

A: Molecule type: mRNA

A: Residues: 1-723 - KAW

A: Cross references: EMBL:U00001; NID:9217979; FIDN:MAA00865.1; PID:9217980

C: Superfamily: histidine ammonia-lyase

C: Keywords: ammonia-lyase, carbon nitrogen lyase

F: 209-211/cross-link: 5 imidazolinone (Ala-gly) *status predicted

F: 210/Modified site: dehydroalanine (Ser) *status predicted

Query Match

Best local similarity 84.2%; Score 3070; DB 2; Length 723;

Matches 594; Conservative 57; Mismatches 51; Indels 6; Gaps 3.

3 NGNINAVVVELICIKPILMGVAALTSILDEVKKVAEPRKPVKLGSEETLVSVGA 62
 21 MAANNMKKVVNS---ADPILMGVAALAEKMSKSHIDPKKMYEPRKPVKLGSEETLVSVGA 77
 63 GIAAANDSPVYVELSEARAGYKASSDWYMSNMKGTDSTGVTGPGATSHRTKOGCA 122
 78 AIAAIDHG--VKVELSEARAGYKASSDWYMSNMKGTDSTGVTGPGATSHRTKOGCA 135
 123 LQKELLRLINAGIFGNGTSHTPSARAAIVRINTLQCYSGIRHEILFATKFLIN 182
 136 LQKELLRLINAGIFGNGTSHTPSARAAIVRINTLQCYSGIRHEILFATKFLIN 195
 183 NNITPELPEKGTITASSGLVPLSYIAGLLTGRPSKAVGPTGEVLAKEAFAAGVEGGE 242
 196 NNITPELPEKGTITASSGLVPLSYIAGLLTGRPSKAVGPTGEVLAKEAFAAGVEGGE 254
 243 PEIQLKRGALAVNGTAVGSMASVIFDANVLAALSEVLSAIFAVYMOCKPRPTDHLTHK 302
 255 FELDPRGEGALVNGTAVGSLASIVLEFANILAVLEVSALFAVYMOCKPRPTDHLTHK 314
 303 LKHHPPQIHAALIMHYIIMQSNYVKAACQVNHMDPIQKPCQRYALPTSPQMLGPQIEV 362
 315 LKHHPPQIHAALIMHYIIMQSNYVKAACQVNHMDPIQKPCQRYALPTSPQMLGPQIEV 374
 363 RSSTKMIETELNSVNDPLIDVSRKALHGNQGTPIGVSMDNTRLATATIGKLMFAOF 422
 375 RPSSTKMIETELNSVNDPLIDVSRKALHGNQGTPIGVSMDNTRLATATIGKLMFAOF 434
 423 SPVNPVYNGCIPSNISGCHNPSIDYCFKQCEIAMAASYSELOFLANVTNIVGSAEON 482
 435 SELVNDFTNNGCLPSNLSASRNSLDYGFGESEIAMAASYSELOFLANVTNIVGSAEON 494
 483 QDVNSLGLISARKTAFAVDLILKMSSTYVALCOSIDLKHLFENKSTYKNTVSOVAKK 542
 495 QDVNSLGLISARKTAFAVDLILKMSSTYVALCOSIDLKHLFENKSTYKNTVSOVAKK 554
 543 LTKMVGSELHPSRCEKDLRYVDREYVAYIDDVCSGTYPLMOKLROYLVDAHINGET 602
 555 LTKMVGSELHPSRCEKDLRYVDREYVAYIDDVCSGTYPLMOKLROYLVDAHINGET 614
 603 EKNNTNISTFOKIAFPPEELKVLIPKFEVPCVRIAYENDTISTNRIKACSYPIYFVREE 662
 615 EKNNTNISTFOKIAFPPEELKVLIPKFEVPCVRIAYENDTISTNRIKACSYPIYFVREE 674
 663 LQKELLRLINAGIFGNGTSHTPSARAAIVRINTLQCYSGIRHEILFATKFLIN 710
 675 LQKELLRLINAGIFGNGTSHTPSARAAIVRINTLQCYSGIRHEILFATKFLIN 722

RESULT 12
 S22991
 Phenylalanine ammonia-lyase (EC 4.3.1.5) 1 - soybean
 C: Species: Glycine max (soybean)
 C: Date: 06-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
 C: Accession: S22991
 R: Frank, R.L.; Wodkin, D.O.
 DNA Seq: 1, 335-346, 1991
 A: Title: Sequence and structure of a phenylalanine ammonia-lyase gene from Glycine max.
 A: Reference number: S22991; MIM: 9219050; PMID: 1794682
 A: Molecule type: DNA

A: Residues: 1-713 - FRA

A: Cross references: EMBL:X52751; NID:918476; FIDN:FAA7129.1; PID:918377

C: Superfamily: histidine ammonia-lyase

C: Keywords: ammonia-lyase, carbon nitrogen lyase

F: 199-201/cross-link: 5 imidazolinone (Ala-gly) *status predicted

F: 200/Modified site: dehydroalanine (Ser) *status predicted

Query Match

Best local similarity 82.6%; Score 1068; DB 2; Length 713;

Matches 592; Conservative 59; Mismatches 54; Indels 12; Gaps 4;

1 MGNQVHNGVNEIC-----KDPINMGVAALTSILDEVKKVAEPRKPVKLGSEETLVSVGA 53
 1 MFATNCHONG---SPCLSTAKCNNDPLINMGAAAFAMKSHIDPEVKMAEPRKPVKLGSEETLVSVGA 57
 54 ETLTVOVAGIAANDSDYKVELSEARAGYKASSDWYMSNMKGTDSTGVTGPGATSHRTKOGCA 113
 58 ETLTVOVAGIAANDSDYKVELSEARAGYKASSDWYMSNMKGTDSTGVTGPGATSHRTKOGCA 115
 114 HRTKOGCALQKELLRLINAGIFGNGTSHTPSARAAIVRINTLQCYSGIRHEILFATKFLIN 173
 116 HRTKOGCALQKELLRLINAGIFGNGTSHTPSARAAIVRINTLQCYSGIRHEILFATKFLIN 175
 174 LEATITPLNNITPCIPPLNGTITASGDIYPLSYIAGLLTGRPSKAVGPTGEVLAKEAFAAGVEGGE 233
 176 LEATITPLNNITPCIPPLNGTITASGDIYPLSYIAGLLTGRPSKAVGPTGEVLAKEAFAAGVEGGE 235
 234 AAGVGFGEFELQPKRGALVNGTAVGSMASVIFDANVLAALSEVLSAIFAVYMOCKPRPTDHLTHK 293
 236 ELASINSRPEYELQPKRGALVNGTAVGSMASVIFDANVLAALSEVLSAIFAVYMOCKPRPTDHLTHK 295
 294 EPTDHLTHKIKHHPPQIHAALIMHYIIMQSNYVKAACQVNHMDPIQKPCQRYALPTSPQMLGPQIEV 353
 296 EPTDHLTHKIKHHPPQIHAALIMHYIIMQSNYVKAACQVNHMDPIQKPCQRYALPTSPQMLGPQIEV 355
 354 WICPQIEVIRFSTKSIEREINSVNDPLIDVSRKALHGNQGTPIGVSMDNTRLATATIGKLMFAOF 413
 356 WICPQIEVIRFSTKSIEREINSVNDPLIDVSRKALHGNQGTPIGVSMDNTRLATATIGKLMFAOF 415
 414 LKLMFAOESSELVNDFTNNGCLPSNLSASRNSLDYGFGESEIAMAASYSELOFLANVTNIVGSAEON 473
 416 LKLMFAOESSELVNDFTNNGCLPSNLSASRNSLDYGFGESEIAMAASYSELOFLANVTNIVGSAEON 475
 474 HVQSAEQHNDVNSLGLISARKTAFAVDLILKMSSTYVALCOSIDLKHLFENKSTYKNTVSOVAKK 533
 476 HVQSAEQHNDVNSLGLISARKTAFAVDLILKMSSTYVALCOSIDLKHLFENKSTYKNTVSOVAKK 535
 534 TVSOVAKKVLITMVGSELHPSRCEKDLRYVDREYVAYIDDVCSGTYPLMOKLROYLVDAHINGET 593
 536 TVSOVAKKVLITMVGSELHPSRCEKDLRYVDREYVAYIDDVCSGTYPLMOKLROYLVDAHINGET 595
 594 DHALNNGTETKNTISTFOKIAFPPEELKVLIPKFEVPCVRIAYENDTISTNRIKACSYPIYFVREE 653
 596 DHALNNGTETKNTISTFOKIAFPPEELKVLIPKFEVPCVRIAYENDTISTNRIKACSYPIYFVREE 655
 654 LTKMVGSELHPSRCEKDLRYVDREYVAYIDDVCSGTYPLMOKLROYLVDAHINGET 710
 656 LTKMVGSELHPSRCEKDLRYVDREYVAYIDDVCSGTYPLMOKLROYLVDAHINGET 712

RESULT 13
 A44133
 Phenylalanine ammonia-lyase (EC 4.3.1.5) - tomato
 C: Species: Lycopersicon esculentum (tomato)
 C: Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 21-May-1999
 C: Accession: A44133; S51076
 R: Lee, S.W.; Kobb, J.; Nazari, R.N.
 J. Biol. Chem. 267, 11824-11830, 1992
 A: Title: Truncated phenylalanine ammonia-lyase expression in tomato (Lycopersicon esc.
 A: Reference number: A44133; MIM: 92291052; PMID: 1601854

Db 679 TGLTGTGNTSPGECDEKLEFAMCGCKITIDPLECEGEMNKAAPLPG 725

RESULT 15
S04127

phenylalanine ammonia-lyase (EC 4.3.1.5) class II - kidney bean

C:Species: Phaseolus vulgaris (kidney bean)

C:Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text-change 21-May-1999

C:Accession: S04127

C:Clamper, C. L.; Edwards, K. Dren, M.; Liang, X.; Hillier, S. G.; Bujewicz, G. P.; Dixon, R.

Plant Mol. Biol. 12: 367-383, 1989

A:Title: Phenylalanine ammonia-lyase gene organization and structure.

A:Reference number: S04127

A:Accession: S04127

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-712 <CRA>

C:Gene: PAL2

C:Gene: PAL2

C:Superfamily: histidine ammonia-lyase

C:Keywords: ammonia-lyase; carbon nitrogen lyase

F:198-200/Cross-link: 5 imidazolinone (Ald-Gly) #status predicted

F:199/Modified site: dehydroalanine (Ser) #status predicted

Query Match 83.9%; Score 3058; DB 2; Length 712;
Best local similarity 84.3%; Pred. No. 3, 9e-188;
Matches 585; Conservative 60; Mismatches 47; Indels 2; Gaps 1;

QY 18 DPLMGVAAEALTGSHLDEYKKVAEERKPVKLGGETLTYSQVAGIAANDSTVVEL 77
Db 21 DPLMMAAAEALSGSHLDEYKRVAAEYRKPAVRLGGQTLIAQVATAHDOG-LKVEL 78
QY 78 SEARACVCAKSSIMVWMSMKGTDSYCVTTGCATSHRRTKOGALOKELIRFLNAGITG 137
Db 79 AESARACVCAKSSIMVWMSMKGTDSYCVTTGCATSHRRTKOGALOKELIRFLNAGITG 138
QY 138 NGTETSHLPKSAIRAMIVIRINTLGGYSGIRFEILEAITFLNNITPCLPLKGTITA 197
Db 139 NGTETSHLPKSAIRAMIVIRINTLGGYSGIRFEILEAITFLNNITPCLPLKGTITA 198
QY 198 SGDIYPLSTYAGLTGTPPNKAVGPTGEVLNAEKAFNAAGVEGPFELQPKGALVNGT 257
Db 199 SGDIYPLSTYAGLTGTPPNKAVGPTGEVLNAEKAFNAAGVEGPFELQPKGALVNGT 258
QY 258 AVSGMASVVPDANVIALISEVLSAIFAVMGKPEFTDHLTHKLKHHQCOIFAAATMR 317
Db 259 AVSGGLASIVLFEENILAVISEVLSAIFAVMGKPEFTDHLTHKLKHHQCOIFAAATMR 318
QY 318 YILDGSDYVAAKVIEMDPLQKPKQDRYALKTSQWLDGPQLEVIKSTKMLEREINSYN 377
Db 319 YILDGSDYVAAKVIEMDPLQKPKQDRYALKTSQWLDGPQLEVIKSTKMLEREINSYN 378
QY 378 DNPLIDYSRKKALHGKNGFSGTPIGVSMONTRLAIAALGKLMAGFSELYNDYNNGLPSN 437
Db 379 DNPLIDYSRKKALHGKNGFSGTPIGVSMONTRLAIAALGKLMAGFSELYNDYNNGLPSN 438
QY 438 LSGGRNPSLDYGFKQCHTAMASYCSHIOPLANFVTHNVOSAPOHNDVNSLQIISARKTA 497
Db 439 LSGGRNPSLDYGFKQCHTAMASYCSHIOPLANFVTHNVOSAPOHNDVNSLQIISARKTA 498
QY 498 EAVDITLKMSTYLVALCGSIDLRHLEENKSTYKNTVSQVAKVLTMGVNGELIIPSRFC 557
Db 499 EAVDITLKMSTYLVALCGSIDLRHLEENKSTYKNTVSQVAKVLTMGVNGELIIPSRFC 558
QY 558 EKDLLRVVDEYVAVYIDVSGSTYPLMOKLROVLVDHALNNGETEKNTNTSIFQKIATF 617
Db 559 EKALLKVEREFTYTFAYTIDDCGSGTYPLMOKLROVLVDHALNNGETEKNTNTSIFQKIATF 618
QY 618 EEEHKKVLLPKRVGQVRIAYVNDITLSIPNRIKACRSYPLVFWVWEIQRGHLTGKVTSPG 677
Db 619 EEEHKKVLLPKRVGQVRIAYVNDITLSIPNRIKACRSYPLVFWVWEIQRGHLTGKVTSPG 678

QY 678 EEEHKKVLLPKRVGQVRIAYVNDITLSIPNRIKACRSYPLVFWVWEIQRGHLTGKVTSPG 711
Db 679 EECDDKVPFAMCGCKITIDPLECEGEMNKAAPLPG 712

Search completed: March 29, 2003, 02:13:15
Job time : 51 secs

.

.

.

.

.

.

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 29, 2003, 00:34:05 : Search time 39 seconds
(without alignments)
756.145 Million cell updates/sec

Title: US-09-964-992A-1
Accession: 3644
Sequence: 1 MENCNHNVCVNVNHC1KIDPL 11DD1LHFC1GCMNCEPLPIC 711

Scoring table: BLAST62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	3206	88.0	667 1	PALY_HELAN
2	3186.5	87.4	716 1	PALY_PORFI
3	3185.5	87.4	716 1	PALY_PORFI
4	3178.5	87.2	716 1	PALY_PORFI
5	3171.5	87.0	708 1	PALY_PORFI
6	3154.5	86.6	714 1	PALY_PORFI
7	3149	86.4	715 1	PALY_PORFI
8	3118.5	85.6	712 1	PALY_PORFI
9	3116	85.5	715 1	PALY_PORFI
10	3115	85.5	717 1	PALY_PORFI
11	3109.5	85.3	712 1	PALY_PORFI
12	3097.5	85.0	721 1	PALY_PORFI
13	3096	85.0	713 1	PALY_PORFI
14	3088	84.7	715 1	PALY_PORFI
15	3077.5	84.5	710 1	PALY_PORFI
16	3077	84.4	722 1	PALY_PORFI
17	3070	84.2	723 1	PALY_PORFI
18	3068.5	84.2	725 1	PALY_PORFI
19	3068	84.2	713 1	PALY_PORFI
20	3067.5	84.2	725 1	PALY_PORFI
21	3058	83.9	712 1	PALY_PORFI
22	3051	83.7	705 1	PALY_PORFI
23	3049	83.7	720 1	PALY_PORFI
24	3049	83.7	724 1	PALY_PORFI
25	3042	83.5	704 1	PALY_PORFI
26	3039.5	83.4	725 1	PALY_PORFI
27	3004	82.4	717 1	PALY_PORFI
28	2994	82.2	718 1	PALY_PORFI
29	2988	82.0	708 1	PALY_PORFI
30	2985.5	81.9	710 1	PALY_PORFI
31	2953.5	81.1	707 1	PALY_PORFI
32	2928.5	80.4	710 1	PALY_PORFI
33	2920.5	80.1	707 1	PALY_PORFI

34	2909	79.8	682 1	PALY_PORFI	p45741 populus kit
35	2908.5	79.8	703 1	PALY_PORFI	q42609 bromelia
36	2730.5	74.9	700 1	PALY_PORFI	q43210 trifolium ae
37	2672.5	73.3	620 1	PALY_PORFI	p43727 trifolium amer
38	2632.5	72.2	710 1	PALY_PORFI	p53443 trifolium amer
39	2590	71.1	710 1	PALY_PORFI	p19143 trifolium amer
40	2585.5	71.0	695 1	PALY_PORFI	p43725 trifolium amer
41	2548	69.6	590 1	PALY_PORFI	p43726 trifolium amer
42	2509.5	68.8	571 1	PALY_PORFI	q40910 populus kit
43	2500	68.6	701 1	PALY_PORFI	p14717 trifolium amer
44	2373.5	65.1	754 1	PALY_PORFI	p52777 trifolium amer
45	2220	60.9	506 1	PALY_PORFI	p52718 trifolium amer

ALIGNMENTS

RESULT 1
ID PALY_HELAN STANDARD: PRI: 667 AA.
AC 004058:
DT 15-JUL-1998 (rel. 36, created)
DT 15-JUL-1998 (rel. 36, last sequence update)
DT 15-JUN-2002 (rel. 41, last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
GN PAL.
OS Helianthus annuus (common sunflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Helianthaceae; Helianthus.
OX NCBI_Taxid:4232;
RN [1]
RP SEQUENCE FROM N.A.
RA Katerina F.A., Salles S., Dreyer J., Brackel-Dreyer P.,
RA Tourville D., Ledoit G.;
RT 'Isolation of a complete pal. cDNA from sunflower.';
RL (in) Plant Gene Register PCR98-108.
CC (in) FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC (in) CATALYTIC ACTIVITY: L-phenylalanine + trans-cinnamate + NH(3).
CC (in) PATHWAY: Phenylpropanoid biosynthesis; first step.
CC (in) SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC (in) PTM: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CIRCULAZION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC (in) SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announc/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: Y124C1, CAA73065.1, ..
CC DR HISPP: P21310; IBBP.
CC DR InterPro: IPR00106, phc/HIS_NH3lyase.
CC DR Pfam: PF00221; PALY_1.
CC DR Trifam: T10601226, phc-am-lyase; 1.
CC DR PROSITE: PS00488; PAL_HISTIDASE; 1.
CC KW Lyase; Phenylpropanoid metabolism; Multiquery family.
CC FT SITE 195 195 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO).
CC FT SITE 195 195 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO).
CC SO SOURCE: 667 AA; 72076 MW; PHD3D590DRH042 CXC64;
CC
CC Query Match: 88.0%, Score 3206, DB 1, Length 667;
CC Best Local Similarity: 94.4%; Pred. No. 1,4e-19;
CC Matches: 627; Conservative: 13; Mismatches: 24; Indels: 0; Gaps: 0;

```

DB 1 MNGGTVNGSANGFCIKDPLMNGVAAEALTSGLDEVKKMGGEPRKKPVVKAQ3ETLTVSQ 60
QY 61 VAGIAAANDSDTVKVELSLAARACKASSDWMHSMKCTDSTVCTTGFCATSHRRKOC 120
DB 61 VAGISAAAGGNGKVELSEAAAGKASSDWMHSMKCTDSTVCTTGFCATSHRRKOC 120
QY 121 GALOKELIFELNAGJFNGTETSHLPHSATKAAMIVRINTLLOGYSIGREILEATKE 180
DB 121 GALOKELIFELNAGJFNGTETSHLPHSATKAAMIVRINTLLOGYSIGREILEATKE 180
QY 181 LNNITPECLPLRGTTTASDVLPSYIAGLLTGPENSKAVPTGEVLNAEKAFAGVEG 240
DB 181 LNNITPECLPLRGTTTASDVLPSYIAGLLTGPENSKAVPTGEVLNAEKAFAGVEG 240
QY 241 GFPELOPKKGLAVNGTAVGSGMASVLEFANVALLSEVLSALFAEVMGKPEFDHLT 300
DB 241 GFPELOPKKGLAVNGTAVGSGMASVLEFANVALLSEVLSALFAEVMGKPEFDHLT 300
QY 301 HKLKHHPGQIEAAAMEYTLDSGVYKAAQVHEMDPLQKPRDRAALRTSPQWLGPOIE 360
DB 301 HKLKHHPGQIEAAAMEYTLDSGVYKAAQVHEMDPLQKPRDRAALRTSPQWLGPOIE 360
QY 361 VTRSSKMLEREINSVNDNPLDVSRRKALHGNFGTPIVSDNTRLAIAIGKLMFA 420
DB 361 VTRSSKMLEREINSVNDNPLDVSRRKALHGNFGTPIVSDNTRLAIAIAIGKLTIA 420
QY 421 QFSELVNDVFNNGPLPSNLSCGNPFLDYGKCEIAMAASYCELOFLANPTNHVQSAEQ 480
DB 421 QFSELVNDVFNNGPLPSNLSCGNPFLDYGKCEIAMAASYCELOFLANPTNHVQSAEQ 480
QY 481 HHQDVNSLGLIARAKTAENVLDIKLMSSTYVALQCSIDLHLEENKSTVKNVQVAK 540
DB 481 HHQDVNSLGLIARAKTAENVLDIKLMSSTYVALQCSIDLHLEENKSTVKNVQVAK 540
QY 541 KYLIMGVNGELHPSRPFCEKDLFVVDPEYFAYITDDVCSGTPTLMOKLROVLVHALNNG 600
DB 541 KYLIMGVNGELHPSRPFCEKDLFVVDPEYFAYITDDVCSGTPTLMOKLROVLVHALNNG 600
QY 601 ETEKNTNTSIFOKIATFPEELKVLPLKEVEGRIAYENDTISIPNRIKACHSYPLAFVVR 660
DB 601 ETEKNTNTSIFOKIATFPEELKVLPLKEVEGRIAYENDTISIPNRIKACHSYPLAFVVR 660
QY 661 EELG 664
DB 661 EELG 664

```

RESULT 2

```

PAL_PETCR STANDARD: PRT: 716 AA.
AC P2481:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
CN PAL.
OS Petroselinum crispum (parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Asterales; euasterids II; Apiales; Apiaceae; Petroselinum.
OX NCBI_TaxID=4043;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=89336635; PubMed=2767049;
RA Lols R., Dietrich A., Hahnbrock K., Schulz W.:
  "A phenylalanine ammonia-lyase gene from parsley: structure,
  regulation and identification of elicitor and light responsive
  cis-acting elements".
  EMBO J. 8:1641-1648(1989).
RN 121
RP SEQUENCE FROM N.A.
RX MEDLINE=89336635; PubMed=2767049;
RA Lols R., Dietrich A., Hahnbrock K., Schulz W.:
  "A phenylalanine ammonia-lyase gene from parsley: structure,
  regulation and identification of elicitor and light responsive
  cis-acting elements".
  EMBO J. 8:1641-1648(1989).
RN 121
RP SEQUENCE FROM N.A.
RX MEDLINE=89336635; PubMed=2767049;
RA Lols R., Dietrich A., Hahnbrock K., Schulz W.:
  "A phenylalanine ammonia-lyase gene from parsley: structure,
  regulation and identification of elicitor and light responsive
  cis-acting elements".
  EMBO J. 8:1641-1648(1989).

```

```

RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP ACTIVE SITE AND MUTAGENESIS.
RX MEDLINE=9432630; PubMed=8050576.
RA Schuster B., Reley J.:
  "Serine-202 is the putative precursor of the active site
  denhydroalanine of phenylalanine ammonia lyase. Site-directed
  mutagenesis studies on the enzyme from parsley (Petroselinum crispum
  L.)".
  FEBS Lett. 349:252-254(1994).
CC -i- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
  FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
  VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -i- CATALYTIC ACTIVITY: L-phenylalanine - trans-cinnamate + NH(3).
CC -i- PATHWAY: phenylpropanoid biosynthesis; first step.
CC -i- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC -i- PFM: CONTAINS AN ACTIVE SITE 4-METHYLDENE-IMIDAZOLE-5-ONE (MIO),
  WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
  OF RESIDUES ALA-SEP-GLY (BY SIMILARITY)
CC -i- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15473; CA33500.1; -.
DR EMBL: X16772; CA34715.1; -.
DR EMBL: Y07654; CA68938.1; -.
DR PIR: S04463; S04463.
DR HSRP: P21310; 1BBP.
DR InterPro: IPR001106; Phe/His_NH3lyase.
DR Pfam: PF00221; PAL. 1.
DR TIGRfams: TIGR01226; pae-am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; 1.
KW Lyase; Phenylpropanoid metabolism; Multigene family.
FT SITE 202 204 MODIFIED TO FORM 4-METHYLDENE-IMIDAZOLE-
  5-ONE (BY SIMILARITY).
FT MUTAGEN 203 203 S->A: COMPLETE LOSS OF ACTIVITY.
FT FT 210 210 S->A: NO LOSS OF ACTIVITY.
SQ SEQUENCE 716 AA; 77628 MW; 23DB2FALACGEBB4A CRC64.

Query Match 87.4%; Score 3186.5; DB 1; Length 716;
Best Local Similarity 86.2%; Pred. No. 2.4e-190;
Matches 620; Conservative 46; Mismatches 42; Indels 11; Gaps 4;

QY 1 MNGGN-----HNGVNVNLCIK--DPLNMGVAAEALTSGLDEVKKMAEPRKPVVKG 52
DB 1 MNGCGATTNGHVNONGMDPCMTEDPLMYGIAEAMTGSGLDEVKMAEVRKPVVKG 60
QY 53 GELLTVSVAGIAAANDSDTVKVELSEAAAGKASSDWMHSMKCTDSTVCTTGFCATSHRRKOC 112
DB 53 GELLTVSVAGIAAANDSDTVKVELSEAAAGKASSDWMHSMKCTDSTVCTTGFCATSHRRKOC 112
QY 61 GETTLISQVAAI-SARDGSGVTVELSEAAAGKASSDWMHSMKCTDSTVCTTGFCATSHRRKOC 119
DB 61 GETTLISQVAAI-SARDGSGVTVELSEAAAGKASSDWMHSMKCTDSTVCTTGFCATSHRRKOC 119
QY 113 SHRRKOCGALOKELIFELNAGJFNGTETSHLPHSATKAAMIVRINTLLOGYSIGREILEATKE 172
DB 113 SHRRKOCGALOKELIFELNAGJFNGTETSHLPHSATKAAMIVRINTLLOGYSIGREILEATKE 172
QY 120 SHRRKOCGALOKELIFELNAGJFNGTETSHLPHSATKAAMIVRINTLLOGYSIGREILEATKE 177
DB 120 SHRRKOCGALOKELIFELNAGJFNGTETSHLPHSATKAAMIVRINTLLOGYSIGREILEATKE 177
QY 173 ILEATKFLNNITPCPLRGTTTASDVLPSYIAGLLTGPENSKAVPTGEVLNAEKAFAGVEG 232
DB 173 ILEATKFLNNITPCPLRGTTTASDVLPSYIAGLLTGPENSKAVPTGEVLNAEKAFAGVEG 232
QY 178 ILEATKFLNNITPCPLRGTTTASDVLPSYIAGLLTGPENSKAVPTGEVLNAEKAFAGVEG 237
DB 178 ILEATKFLNNITPCPLRGTTTASDVLPSYIAGLLTGPENSKAVPTGEVLNAEKAFAGVEG 237
QY 233 FAAGVEGFELOPKKGLAVNGTAVGSGMASVLEFANVALLSEVLSALFAEVMGKPEFDHLT 292
DB 233 FAAGVEGFELOPKKGLAVNGTAVGSGMASVLEFANVALLSEVLSALFAEVMGKPEFDHLT 292
QY 238 FKLAVGEGFELOPKKGLAVNGTAVGSGMASVLEFANVALLSEVLSALFAEVMGKPEFDHLT 297
DB 238 FKLAVGEGFELOPKKGLAVNGTAVGSGMASVLEFANVALLSEVLSALFAEVMGKPEFDHLT 297
QY 293 PEFDHLTTHLKHHPGQIEAAAMEYTLDSGVYKAAQVHEMDPLQKPRDRAALRTSP 352
DB 293 PEFDHLTTHLKHHPGQIEAAAMEYTLDSGVYKAAQVHEMDPLQKPRDRAALRTSP 352

```

OY 353 QMIGPOLFVIRSTKMIEREINSVNDNPLIDVSRNKAJHGNFOCTPIGVSMNDTRIALA 412
 DB 358 QMIGPOLFVIRSTKMIEREINSVNDNPLIDVSRNKAJHGNFOCTPIGVSMNDTRIALA 417
 OY 413 AICKLMFAQSELVNDYNNGLPSNLSAHRNPSLDYVSKAEIAMAASVSEIOPIANVT 472
 DB 418 AICKLMFAQSELVNDYNNGLPSNLSAHRNPSLDYVSKAEIAMAASVSEIOPIANVT 477
 OY 473 NHVQSAEJHNDVNSLGLISAKRTAEVDILKMSSTYLVALCOSIDIRHEENKSTVK 532
 DB 478 NHVQSAEJHNDVNSLGLISAKRTAEVDILKMSSTYLVALCOSIDIRHEENKSTVK 537
 OY 533 NHVQSAEJHNDVNSLGLISAKRTAEVDILKMSSTYLVALCOSIDIRHEENKSTVK 592
 DB 538 NHVQSAEJHNDVNSLGLISAKRTAEVDILKMSSTYLVALCOSIDIRHEENKSTVK 597
 OY 593 VDHALNNGEFTKNTNIFOKIATFEELKVLLEKVEGVKRAYENDLISIPNRKAKRS 652
 DB 598 VDHALNNGEFTKNTNIFOKIATFEELKVLLEKVEGVKRAYENDLISIPNRKAKRS 657
 OY 653 YPLKFEVREELKVEFTGKVTSPGEFVFTAMKRIIDPLELJLJHNMGEPLPIC 711
 DB 658 YPLKFEVREELKVEFTGKVTSPGEFVFTAMKRIIDPLELJLJHNMGEPLPIC 716
 RESULT 3
 PAL3_PETCR STANDARD: PRT: 716 AA.
 AC P45728;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Phenylalanine ammonia-lyase 2 (EC 4.3.1.5).
 GN PAL2.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota: Viridiplantae, Streptophyta: Embryophyta, Tracheophyta;
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
 OC Asterales: Cuscutridae: Apiales: Apiales: Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95010141; PubMed=7925471;
 RA Appert C., Logemann K., Halbrock K., Schmid J., Amrhein N.;
 RT Structural and catalytic properties of the four phenylalanine
 RT ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.);
 PL Eur J Biochem 225:491-499(1994).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine - trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Prokaryote).
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLHISTIDINE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDINASE FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X81158, CAA57056.1;
 DR HSSP: P21310; 188F;
 DR InterPro: IPR001106; Phc/His_NH3lyase.
 DR Pfam: PF00221; PAL_1;
 DR TIGRfam: TIGR01226; Phc-am_lyase; 1;
 DR PROSITE: PS00484; PAL_HISTIDASE; 1;
 KW Lyase, Phenylpropanoid metabolism, Multigene family.
 FT SITE 202 204 MODIFIED TO FORM 4-METHYLHISTIDINE-IMIDAZOLE-5-ONE.

FT 5-ONE (BY SIMILARITY).
 SQ SEQUENCE: 716 AA: 77964 MW: 17470881.88666 CMO64;
 Query Match 87.4%; Score 3185.5; DB 1: Length 716;
 Best Local Similarity 86.2%; Fred No 2 Me-140;
 Matches 620; Conservative 45; Mismatches 43; Indels 11; Gaps 4;
 OY 1 MEMCN-----HNGCVNNEICIK--DPIIMGVAALATGSHIDYKKKVAAPRRPVKIG 52
 DB 1 MEMCNATITNGHNGVNGKDFCKTEDEPLTWGTAAPAMTSHLFEKKKVAAPRRPVKIG 60
 OY 53 GETHLVQVAGIAAAMIDIVKVELLSFAKAGVAKASSHWVMVSMKKGIISVGVIGQVAT 112
 DB 61 GETHLVQVAAI-SADGSGVITVELLSFAKAGVAKASSHWVMVSMKKGIISVGVIGQVAT 119
 OY 113 SHRTKQGNALQKELTLPNAGIFPANGTETSITLPHSTPAAMIVITLQVSGITFE 172
 DB 120 SHRTKQGNALQKELTLPNAGIFPANGTETSITLPHSTPAAMIVITLQVSGITFE 177
 OY 173 ILBATKPLNNNTPLCPPLKGTITASGDIPLISYLAGITGCPNSKAVGPIGVINAKPA 232
 DB 178 ILBATKPLNNNTPLCPPLKGTITASGDIPLISYLAGITGCPNSKAVGPIGVINAKPA 237
 OY 233 PAAQVQCGFPFLQPKGLAVNGTAVGSGMASMVLFPANVIALSEVLSAIFARVMCK 292
 DB 238 PAAQVQCGFPFLQPKGLAVNGTAVGSGMASMVLFPANVIALSEVLSAIFARVMCK 297
 OY 293 PFTDILTIKRLKHPVUTGAAIMETLDSSTVYAAQKVIEMDPLQPKGTPVALPSTP 352
 DB 298 PFTDILTIKRLKHPVUTGAAIMETLDSSTVYAAQKVIEMDPLQPKGTPVALPSTP 357
 OY 353 QMIGPOLFVIRSTKMIEREINSVNDNPLIDVSRNKAJHGNFOCTPIGVSMNDTRIALA 412
 DB 358 QMIGPOLFVIRSTKMIEREINSVNDNPLIDVSRNKAJHGNFOCTPIGVSMNDTRIALA 417
 OY 413 AICKLMFAQSELVNDYNNGLPSNLSAHRNPSLDYVSKAEIAMAASVSEIOPIANVT 472
 DB 418 AICKLMFAQSELVNDYNNGLPSNLSAHRNPSLDYVSKAEIAMAASVSEIOPIANVT 477
 OY 473 NHVQSAEJHNDVNSLGLISAKRTAEVDILKMSSTYLVALCOSIDIRHEENKSTVK 532
 DB 478 NHVQSAEJHNDVNSLGLISAKRTAEVDILKMSSTYLVALCOSIDIRHEENKSTVK 537
 OY 533 NHVQSAEJHNDVNSLGLISAKRTAEVDILKMSSTYLVALCOSIDIRHEENKSTVK 592
 DB 538 NHVQSAEJHNDVNSLGLISAKRTAEVDILKMSSTYLVALCOSIDIRHEENKSTVK 597
 OY 593 VDHALNNGEFTKNTNIFOKIATFEELKVLLEKVEGVKRAYENDLISIPNRKAKRS 652
 DB 598 VDHALNNGEFTKNTNIFOKIATFEELKVLLEKVEGVKRAYENDLISIPNRKAKRS 657
 OY 653 YPLKFEVREELKVEFTGKVTSPGEFVFTAMKRIIDPLELJLJHNMGEPLPIC 711
 DB 658 YPLKFEVREELKVEFTGKVTSPGEFVFTAMKRIIDPLELJLJHNMGEPLPIC 716
 RESULT 4
 PAL3_PETCR STANDARD: PRT: 716 AA.
 AC P45729;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Phenylalanine ammonia-lyase 3 (EC 4.3.1.5).
 GN PAL3.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota: Viridiplantae, Streptophyta: Embryophyta, Tracheophyta;
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
 OC Asterales: Cuscutridae: Apiales: Apiales: Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95010141; PubMed=7925471;

RA Appert C., Logemann E., Hahlbrock K., Schmid J., Amrhein N.;
 RT "Structural and catalytic properties of the four phenylalanine
 RT ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.).";
 RL Eur. J. Biochem. 225:491-499(1994).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis: first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- Ptm: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X81159; CA57057.1; -.
 DR HSPF: P21310; 188F.
 DR InterPro: IPR001106; Phe/His_NH3lyase.
 DR Pfam: PF00221; PAL; 1.
 DR TIGRFAMS: TIGR01226; phe.am.lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE; 1.
 KM Lyase; Phenylpropanoid metabolism; Multigene family.
 FT SITE 204 206 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-
 FT 5-ONE (BY SIMILARITY).
 SQ SEQUENCE 718 AA; 78165 MW; 17786451FFBDC35 CRC64;
 Query Match 87.2%; Score 3178.5; DB 1; Length 718;
 Best Local Similarity 87.1%; Pred. No. 77e-190;
 Matches 620; Conservative 46; Mismatches 41; Indels 5; Gaps 4;
 3 NGNHVGVNVELCIK-DPLMGVAEALTGSHDEYKKNVAERKRPVYKLGTELTYSQ 60
 9 NC-HANGNDIDLCMKREPLMNGVAEALTGSHDEVYRMVAEYKRPVYKLGTELTISQ 67
 61 VAGIAANDSDTYKVELSEARAGYKASSDVWMSNMKGTSYCVTTGFGATSHRRKQG 120
 68 VAAISKRDSG-VKVELSEARAGYKASSDVWMSNMKGTSYCVTTGFGATSHRRKQG 126
 121 GALQKELIRFLNAGIFGNCQTF-SHTLPHSATRAAMIVRINTLQGYSGIRFELLEATIK 179
 127 GALQKELIRFLNAGIFGSGAENGNTLPHSATRAAMIVRINTLQGYSGIRFELLEATIK 186
 180 PLANNITTCPLRGHTTASGDLVPLSYTAGLLTGRPNKAVGPTGEVLAERKFAAGVE 239
 187 PLNNHTTPLPLRGHTTASGDLVPLSYTAGLLTGRPNKAVGPTGEVLAERKFAAGVE 246
 240 GGEFFELQPEGLATNGTAVSGMAWYLPDANVALLSEVLSAIFAEVMOGKREFPDHL 299
 247 GGEFFELQPEGLATNGTAVSGMAWYLPDANVALLSEVLSAIFAEVMOGKREFPDHL 306
 300 TKLKHHPQQLFAAAIMELYILDGSDYVAAQKVHEMDPLQPKODRYALFTSPQWLQPOL 359
 307 TKLKHHPQQLFAAAIMELYILDGSAVYKAAKLHMDPLQPKODRYALFTSPQWLQPOL 366
 360 EYIRSTKMIEREINSVNDNPLIDVSRNKALHGNFCQTPRGVSMNDTRLAIAIGKLMF 419
 367 EYIRSTKMIEREINSVNDNPLIDVSRNKALHGNFCQTPRGVSMNDTRLAIAIGKLMF 426
 420 AOFSELYNDEVYNGJPSNLSGGRNSLDYGFKEGELIAMSVCSELOFLANVYNHQSAAE 479
 427 AOFSELYNDEVYNGJPSNLSGGRNSLDYGFKEGELIAMSVCSELOFLANVYNHQSAAE 486
 480 QHNDVNSLGLISARKTEAVDILKLMSSYTLVALCOSIDRLHLEEMKSTVKTYSOVA 539
 487 QHNDVNSLGLISARKTEAVDILKLMSSYTLVALCOSIDRLHLEEMKSTVKTYSOVA 546

QY 540 KVVILMGVNGELHPSRFECKDILLRVDRVEYFAVIDVCGSTYPLMOKLROVYDHALNN 599
 DB 547 KVVILMGVNGELHPSRFECKDILLRVDRVEYFAVIDVCGSATYPLMOKLRETLVEHALNN 606
 QY 600 GFEKNTSTISFOKATATEEELKYLPEVEGVIAVENLTSPNIRKACRSYPLRYEF 659
 DB 607 GKKNELSTISFOKIAAEDEELKYLPEVEGTAAALSNPAINPKRCRSYPLRYEF 666
 QY 660 REELRGELTGKVTSPGEFDRVPTAMCKQIIDLPLECLGNGHPDIC 711
 DB 667 RELGTETVLTGKRVSPGEFEKVTAMSKKEIIDPLECLGNGAPDIC 718
 RESULT 5
 ID PALI_DAUCA STANDARD; PRT; 708 AA.
 AC 023865;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
 GN PALI.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Kurodagosun;
 RX MEDLINE=97478956; PubMed=9337617;
 RA Takeda J., Ozeki Y., Yoshida K.;
 RT "Action spectrum for induction of promoter activity of phenylalanine
 RT ammonia-lyase gene by UV in carrot suspension cells.";
 RL Photochem. Photobiol. 66:464-470(1997).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis: first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- Ptm: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D85850; BAA23367.1; -.
 DR HSPF: P21310; 188F.
 DR InterPro: IPR001106; Phe/His_NH3lyase.
 DR Pfam: PF00221; PAL; 1.
 DR TIGRFAMS: TIGR01226; phe.am.lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE; 1.
 KM Lyase; Phenylpropanoid metabolism; Multigene family.
 FT SITE 194 196 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-
 FT 5-ONE (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76845 MW; 0610411373E680E3 CRC64;
 Query Match 87.0%; Score 3171.5; DB 1; Length 708;
 Best Local Similarity 87.2%; Pred. No. 2e-189;
 Matches 619; Conservative 41; Mismatches 43; Indels 7; Gaps 3;
 2 ENGNHVNVEVLECI-KQPLMGVAEALTGSHDEYKKNVAERKRPVYKLGTELTYSQ 60
 DB 4 ENKNVYVLG-NGLCMKQDPLMNGVAEALTGSHDEYKKNVAERKRPVYKLGTELTYSQ 61

QY 61 VAAIAAAG----SVKVLHSAARAGVASSDWMWESNNKKTDSYVITGATSHHRTKOG 120
 Db 62 VAAIAAAG----SVKVLHSAARAGVASSDWMWESNNKKTDSYVITGATSHHRTKOG 117
 QY 121 GALLKELIRFLNAGIHPNTESSHTLPHSATRAAMIVPINTLLGYSGSIPEILEATKE 180
 Db 118 GALLKELIRFLNAGIHPNTESSHTLPHSATRAAMIVPINTLLGYSGSIPEILEATKE 177
 QY 181 LNNVITPCPLRGCTTASGDIAPLSTAGLITGRPNKAVGPTGVNAERKFAAGAG 240
 Db 178 LNNVITPCPLRGCTTASGDIAPLSTAGLITGRPNKAVGPTGVNAERKFAAGAG 237
 QY 241 GFEELQKREGIALVNGTAVSGSMASVYLDANVALLSEVLSAIFAEVQKPEFTDILT 300
 Db 238 GFEELQKREGIALVNGTAVSGSMASVYLDANVALLSEVLSAIFAEVQKPEFTDILT 297
 QY 301 HKLKHHPGUEAAAMEYILDSSDYKAAAKVHEMPPLKPKUDRYALFTSPQWLSPOLE 360
 Db 298 HKLKHHPGUEAAAMEYILDSSDYKAAAKVHEMPPLKPKUDRYALFTSPQWLSPOLE 357
 QY 361 VHSSTKMIEREINSVNDNPLDVSFNKALHGNFQCPPTIGVSMNTPLAIAIGKLMPA 420
 Db 358 VHSSTKMIEREINSVNDNPLDVSFNKALHGNFQCPPTIGVSMNTPLAIAIGKLMPA 417
 QY 421 QFSELVNFYNNLPLSNLSAGRNPSLDYGRKSSIELAMASYCSELPANPVTNHVSADQ 480
 Db 418 QFSELVNFYNNLPLSNLSAGRNPSLDYGRKSSIELAMASYCSELPANPVTNHVSADQ 477
 QY 481 HNDVNSLGLISARKTAENDTLKLSSTYVLAQSTDLRLLEENKSTVKNYTSQAK 540
 Db 478 HNDVNSLGLISARKTAENDTLKLSSTYVLAQSTDLRLLEENKSTVKNYTSQAK 537
 QY 541 KVLTMGVNGLHPSRCKFKDIAKVDRPFVFAVTDVCSGTPLMKLIKQVLDVDAIANN 600
 Db 538 KVLTMGVNGLHPSRCKFKDIAKVDRPFVFAVTDVCSGTPLMKLIKQVLDVDAIANN 597
 QY 601 ETEKNTSTFQKATFEEELKVLKREYEGVPIAVENDTLSTPNRKAQSYPLYPVR 660
 Db 598 ETEKNTSTFQKATFEEELKVLKREYEGVPIAVENDTLSTPNRKAQSYPLYPVR 657
 QY 661 EELGRGLTGEKVTSPGEEDRYETAMCGOIIDPLELGLGMMNCPDPI 710
 Db 658 EELGRGLTGEKVTSPGEEDRYETAMCGOIIDPLELGLGMMNCPDPI 707

RESULT 6
 ID PALY CAMSI STANDARD: PRI: 714 AA.
 AC P45726;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
 OS PAL.
 OS Camellia sinensis (Tea)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 OC Asteridaceae; Ericaceae; Theaceae; Camellia.
 OX NCBI_TaxID=4442;
 RN [1]
 RP SEQUENCE FROM N.A.
 KC STRAIN=CV. Yabukita; TISSUE=leaf;
 RA Matsumoto S., Takuchi A., Hayatsu M., Kondo S.;
 RT Molecular cloning of phenylalanine ammonia-lyase cDNA and
 RT classification of varieties and cultivars of tea plants (Camellia
 RT sinensis) using the tea PAL cDNA probe.*
 RL Theor. Appl. Genet. 89:671-675(1994)
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine -> trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC
 CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC
 CC FMRI: D26596; RAA05643.1;
 CC HSSP: P21310; 188F.
 CC InterPro: IPR001106; Phe/His_NH3lyase.
 CC Pfam: PF00221; PAL.1.
 CC TIGRFAMs: TIGR01226; phe_am_lyase.1.
 CC PROSITE: PS00488; PAL_HISTIDASE.1.
 CC Lysase; Phenylpropanoid metabolism.
 CC SITE 200 202 MODIFIED TO FORM 4 METHYLENE-IMIDAZOLE-5-ONE (BY SIMILARITY).
 CC SEQUENCE 714 AA: 77751 MW: 50607.800 Da
 CC
 CC Query Match 86.6%, Score 3154.5, DB 1, Length 714;
 CC Best Local Similarity 86.0%; Pred. No. 2,40-188;
 CC Matches 610; Conservative 46; Mismatches 50; Indels 3; Gaps 2;

QY 4 GHHV-NCVNEELCKDPLNKGVAEALTSGLDEKKNVAERKRVKVGSETLTVSQA 62
 Db 8 GHHV-NCVNEELCKDPLNKGVAEALTSGLDEKKNVAERKRVKVGSETLTVSQA 67
 QY 63 GAAANSPDVVKKELSPAAKAGVASSDWMWESNNKKTDSYVITGATSHHRTKOG 122
 Db 62 GAAANSPDVVKKELSPAAKAGVASSDWMWESNNKKTDSYVITGATSHHRTKOG 119
 QY 68 AIAVAGSE--VAVELSESAREVKAASSDWMWESNNKKTDSYVITGATSHHRTKOG 125
 Db 67 AIAVAGSE--VAVELSESAREVKAASSDWMWESNNKKTDSYVITGATSHHRTKOG 122
 QY 123 LQRELFIRFLNAGIHPNTESSHTLPHSATRAAMIVPINTLLGYSGSIPEILEATKE 182
 Db 126 LQRELFIRFLNAGIHPNTESSHTLPHSATRAAMIVPINTLLGYSGSIPEILEATKE 185
 QY 183 NNTTTCPLRGCTTASGDIAPLSTAGLITGRPNKAVGPTGVNAERKFAAGAG 240
 Db 186 NNTTTCPLRGCTTASGDIAPLSTAGLITGRPNKAVGPTGVNAERKFAAGAG 237
 QY 243 GFEELQKREGIALVNGTAVSGSMASVYLDANVALLSEVLSAIFAEVQKPEFTDILT 300
 Db 246 GFEELQKREGIALVNGTAVSGSMASVYLDANVALLSEVLSAIFAEVQKPEFTDILT 297
 QY 303 LKHHPGUEAAAMEYILDSSDYKAAAKVHEMPPLKPKUDRYALFTSPQWLSPOLE 360
 Db 306 LKHHPGUEAAAMEYILDSSDYKAAAKVHEMPPLKPKUDRYALFTSPQWLSPOLE 357
 QY 363 VHSSTKMIEREINSVNDNPLDVSFNKALHGNFQCPPTIGVSMNTPLAIAIGKLMPA 420
 Db 366 VHSSTKMIEREINSVNDNPLDVSFNKALHGNFQCPPTIGVSMNTPLAIAIGKLMPA 417
 QY 421 QFSELVNFYNNLPLSNLSAGRNPSLDYGRKSSIELAMASYCSELPANPVTNHVSADQ 480
 Db 423 QFSELVNFYNNLPLSNLSAGRNPSLDYGRKSSIELAMASYCSELPANPVTNHVSADQ 419
 QY 481 HNDVNSLGLISARKTAENDTLKLSSTYVLAQSTDLRLLEENKSTVKNYTSQAK 540
 Db 486 HNDVNSLGLISARKTAENDTLKLSSTYVLAQSTDLRLLEENKSTVKNYTSQAK 537
 QY 541 KVLTMGVNGLHPSRCKFKDIAKVDRPFVFAVTDVCSGTPLMKLIKQVLDVDAIANN 600
 Db 543 KVLTMGVNGLHPSRCKFKDIAKVDRPFVFAVTDVCSGTPLMKLIKQVLDVDAIANN 599
 QY 601 ETEKNTSTFQKATFEEELKVLKREYEGVPIAVENDTLSTPNRKAQSYPLYPVR 660
 Db 603 ETEKNTSTFQKATFEEELKVLKREYEGVPIAVENDTLSTPNRKAQSYPLYPVR 657
 QY 661 EELGRGLTGEKVTSPGEEDRYETAMCGOIIDPLELGLGMMNCPDPI 710
 Db 606 EELGRGLTGEKVTSPGEEDRYETAMCGOIIDPLELGLGMMNCPDPI 707

663 LKHS:FLTS:PKVTS:SEEFDFPTAM:K:K:IDPILLETG:SWN:EPITC 711
 DB 666 JOTELITG:HKVNS:PGSEDFKVTAL:CKGEM:IDPLMDCKEM:GAP:PIC 714

RESULT 7
 PAL3_TOBAC STANDARD: PRT: 715 AA.

AC P25872:
 DB 01-MAY-1992 (Rel. 22, Created)
 DB 01-MAY-1992 (Rel. 22, Last sequence update)
 DB 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
 GN TPAL OR PALB.
 OS Nicotiana tabacum (Common tobacco).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 CC NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fukasawa-Akeda T.;
 RI Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
 FN [2]
 RP SEQUENCE FROM N.A
 RC STRAIN=cv. Bright yellow; TISSUE=Callus;
 RA Taguchi G., Sharan M., Gonda K., Yanagisawa K., Shimozaki M.,
 RA Hayashida N., Okazaki M.;
 RI Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine - trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis: first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M84466; AAA34122.1;
 DR EMBL: AB008200; BAA22948.1;
 DR HSPF: P21310; 188F.
 DR HSPF: P21310; 188F.
 DR InterPro: IPR001106; Phe/HIS_NH3lyase.
 DR Pfam: PF00221; PAL; 1.
 DR TIGRFAMs: TIGR01226; phe_am_lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE; 1.
 KM Lyase; Phenylpropanoid metabolism; Multigene family.
 FT SITE 201 203 MODIFIED TO FORM 4-METHYLENE-IMIDAZOLE-5-ONE (BY SIMILARITY).
 FT SITE 201 203 MODIFIED TO FORM 4-METHYLENE-IMIDAZOLE-5-ONE (BY SIMILARITY).
 SQ SEQUENCE 715 AA: 77780 MW: 03GBAE8527394C62 CKC64;

Query Match 86.4%; Score 3149; DB 1; Length 715;
 Best Local Similarity 85.3%; Pred. No. 5.2e-188;
 Matches 610; Conservative 45; Mismatches 56; Indels 4; Gaps 3;

QY 1 MENGHNVCVN-ELCIR--DPIANGVAFAI:GSHDEKVMVAEPKPVKLGRTIT 57
 DB 1 MASNIGNVGENFELCKKSADPLNEMAAESIRGSHDEKVMVAEPKPVKLGRTIT 60
 QY 58 VQVQVAAIANS-DTAVVEISEARAGVAKSSDMVMSNMKGTIVSGYTTGFPATSHR 116
 DB 61 VAQVAAIAVRDSANGVAVKELSEARAGVAKSSDMVMSNMKGTIVSGYTTGFPATSHR 120
 QY 117 TKQSGALQKELIFLNIAGIFGNTETSHLPHSATRAMIVFNTLQVYSGIFELTEA 176

DB 121 TKNQALIDKELIRPLNAVFPNGTETSHLPHSATRAMIVFNTLQVYSGIFELTEA 180
 QY 177 ITRPLNNITPCLPDKCTTASGDLVPLVYIAGILTPRNSKAVGPTGVINAEPKFAAA 236
 DB 181 ITRKINSNTPLPLPGTTASGDLVPLVYIAGILTPRNSKAVGPTGVINAEPKFAAA 240
 QY 237 GVEGPFELUPKEGALVNGTAVGSGMASWYLPANLALLSEVLSIAFEVNOGPEET 296
 DB 241 GVGSGPFELUPKEGALVNGTAVGSGMASWYLPANLALLSEVLSIAFEVNOGPEET 300
 QY 297 DDLTKLKHKKVQLEAAAMEYILDSDYKAAAKVHEMDPLKFKQDFALRTSFQMG 356
 DB 301 DDLTKLKHKKVQLEAAAMEYILDSDYKAAAKVHEMDPLKFKQDFALRTSFQMG 360
 QY 357 POIEVIRSTKMIEREINSVNDPLIDVSRKKAHGCNPGCTPIGVSMDTRIALAIGK 416
 DB 361 POIEVIRSTKMIEREINSVNDPLIDVSRKKAHGCNPGCTPIGVSMDTRIALAIGK 420
 QY 417 LMEAFSELYNDEYNGSLPSNLGSRNPSLDYGEKGEIIMASVSELOFLANPTNHYO 476
 DB 421 LMEAFSELYNDEYNGSLPSNLGSRNPSLDYGEKGEIIMASVSELOFLANPTNHYO 480
 QY 477 SAEJHNDVNSLGLISAKRTAEAVDILKMSSTYLVALQSIDPLHEENKSTVNTVS 536
 DB 481 SAEJHNDVNSLGLISAKRTAEAVDILKMSSTYLVALQSIDPLHEENKSTVNTVS 540
 QY 537 QVAKVITMGVNGELHPRCEKDLRVDPREYFAYITDVGSTYPIMOKI:ROVIVDHA 596
 DB 541 QVAKVITMGVNGELHPRCEKDLRVDPREYFAYITDVGSTYPIMOKI:ROVIVDHA 600
 QY 597 LNNGETENNTSIFOKATFEELKVLPRKEVGVAIVENDTSLIPNKIKACRSYPL 656
 DB 601 MNGSEKKNVNSIIPQKIGAREDELKAVLPKEVSAARAALESQNALPNIITCRSPLY 660
 QY 657 KVPPELHGELTSEKVTISGEEDFPVPTAM:K:K:IDPILLETG:SWN:EPITC 711
 DB 661 KVPPELHGELTSEKVTISGEEDFPVPTAM:K:K:IDPILLETG:SWN:EPITC 715

RESULT 8
 PAL3_TOBAC STANDARD: PRT: 712 AA.

AC P45733:
 DB 01-NOV-1995 (Rel. 32, Created)
 DB 01-NOV-1995 (Rel. 32, Last sequence update)
 DB 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
 GN Nicotiana tabacum (Common tobacco).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 CC NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=cv. Samsun NN; TISSUE=leaf;
 RA MEDLINE=95125127; PubMed=7824656;
 RA Pellegrini L., Kohfutsch O., Fritig B., Legend M.;
 RI expression during the hypersensitive reaction to tobacco mosaic virus
 RI and the response to a fungal elicitor.*;
 RT Plant Physiol. 106:877-886(1994).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine - trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis: first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- INDUCTION: STRONGLY INDUCED DURING THE HYPERSENSITIVE REACTION TO
 CC TMV OR TO A FUNGAL ELICITOR.
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: X78269; CAA55075.1; .

DR HSSP: P21310; 1B8F.

DR InterPro: IPR001106; Phe/His_NH3lyase

DR Pfam: PF00221; PAL; 1.

DR TIGRPFAM: TIGR01226; Phe-am_lyase; 1.

DR PROSITE: PS00488; PAL_HISTIDASE; 1.

KM Lyase, Phenylpropanoid metabolism, Multigene family.

FT SITE 198 200 MODIFIED TO FORM 4-METHYLDENE-IMIDAZOLE-5-ONE (BY SIMILARITY).

FT SITE 198 200

SQ SEQUENCE 712 AA: 77510 MW: 66585FA0A83244F0 CRC64:

Query Match: 85.6%, Score 3118.5, DB 1, Length 712;

Host Local Similarity 86.0%; Pred. No. 4e-186;

Matches 602; Conservative 38; Mismatches 59; Indels 1; Gaps 1;

QY 13 ELGK-DPLNMGVAALITSHLEDEYKKNVAFRRKPVYVLLSGELTSSQVAGIAANDSD 71

DB 13 DPCVKVPIINMHEMAADSLGSHIDHYKKNVAFRRKPVYVLLSGELTSSQVAGIAANDSD 72

QY 72 TVVEVLESEARACVAKASVWMSMKSTDSYVTTGATSHPTKQYALQKELIFPL 131

DB 73 TVVEVLESEARACVAKASVWMSMKSTDSYVTTGATSHPTKQYALQKELIFPL 132

QY 132 NAGIFGNGTSTHPLHSATRAAMIVRNILQSGYSGIFELLHATKRLANNITPCLPL 191

DB 133 NAGIFGNGTSTHPLHSATRAAMIVRNILQSGYSGIFELLHATKRLANNITPCLPL 192

QY 192 RGTITASGDLVPLSYLA:LLTTRPNSKAV:PT:EVLLNAERFAAGVGGFELQPKESL 251

DB 193 RGTITASGDLVPLSYLA:LLTTRPNSKAV:PT:EVLLNAERFAAGVGGFELQPKESL 252

QY 252 ALVNGTAVSSGMAVYLFQANVLLSEVLSAFAVYVQAGKEFDLTHIKKHPGLE 311

DB 253 ALVNGTAVSSGMAVYLFQANVLLSEVLSAFAVYVQAGKEFDLTHIKKHPGLE 312

QY 312 AAALMEVYLDGSDYVKAAGVHMDPLQKIKQIKYALKISQWMLGHOIFVIRSSTKMIER 371

DB 313 AAALMEVYLDGSDYVKAAGVHMDPLQKIKQIKYALKISQWMLGHOIFVIRSSTKMIER 372

QY 372 EINSVNDNFLIDVSRKKALHGN:FGQTPITGVSMNDTRIALTAATGKLMFAQSELYNDEVN 431

DB 373 EINSVNDNFLIDVSRKKALHGN:FGQTPITGVSMNDTRIALTAATGKLMFAQSELYNDEVN 432

QY 432 NGIPSNISGGRRNSLDYGRKGSFANASYSCELOFLANVTNIVOSADQINDVSGIT 491

DB 433 NGIPSNISGGRRNSLDYGRKGSFANASYSCELOFLANVTNIVOSADQINDVSGIT 492

QY 492 SAKRTAAVNI::KIMSTYIVAI:QSLDINHL:ENNKSTVKKNIYQVAAKV:ILMVGNDL 551

DB 493 SAKRTAAVNI::KIMSTYIVAI:QSLDINHL:ENNKSTVKKNIYQVAAKV:ILMVGNDL 552

QY 552 IPRSFCEKELIARVYDEYFAVYTDVGS:TYPLMOKLROYLVLDLNNETFRKNTSTIF 611

DB 553 IPRSFCEKELIARVYDEYFAVYTDVGS:TYPLMOKLROYLVLDLNNETFRKNTSTIF 612

QY 612 OKITAFEEELKVLPEVEVSVRIAYENDTLSIPNRIKARSPVCLRFVEVEELGPRFTIGE 671

DB 613 OKITAFEEELKVLPEVEVSVRIAYENDTLSIPNRIKARSPVCLRFVEVEELGPRFTIGE 672

QY 672 KVISPGHEHEDVATAMCKGOIIMYLLKCIQGNNGRP:PLIC 711

DB 673 KVISPGHEHEDVATAMCKGOIIMYLLKCIQGNNGRP:PLIC 712

RESULT 9

ID PALY_POPTR STANAKU: PRT: 715 AA.

AC 945730;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-NOV-2002 (Rel. 41, Last annotation update)

DE Phenylalanine ammonia-lyase (EC 4.3.1.5).

OS PAL.

ON Populus trichocarpa (Western balsam poplar).

CC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;

CC Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae;

CC eurosids I: Malpighiales, Salicaceae, Populus

CC NCBI_TaxID=3694;

RN 111

RP SEQUENCE FROM N.A.

RC STRAIN=P. trichocarpa X P. deltoides: TISSUE=leaf;

FX MEDLINE=94151434; PubMed=4108506;

FA Substratum R. Reinold's, Mollitor E.K., Douglas C.J.;

KT Structure, inheritance, and expression of hybrid poplar (Populus trichocarpa x Populus deltoides) phenylalanine ammonia-lyase genes.*;

PL Plant Physiol. 102:71-83(1993).

CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.

CC -1- CATALYTIC ACTIVITY: L-phenylalanine -> trans-cinnamate + NH3.

CC -1- PATHWAY: Phenylpropanoid biosynthesis, first step.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLDENE-IMIDAZOLE-5-ONE (MIO), WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION OF RESIDUES ALA-SER-GLY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: L11747; AAA33805.1; .

DR HSSP: P21310; 1B8F.

DR InterPro: IPR001106; Phe/His_NH3lyase.

DR Pfam: PF00221; PAL; 1.

DR TIGRPFAM: TIGR01226; Phe-am_lyase; 1.

DR PROSITE: PS00488; PAL_HISTIDASE; 1.

KM Lyase, Phenylpropanoid metabolism, Multigene family.

FT SITE 201 203 MODIFIED TO FORM 4-METHYLDENE-IMIDAZOLE-5-ONE (BY SIMILARITY).

FT SITE 201 203

SQ SEQUENCE 715 AA: 77918 MW: 5896837AB8AA246 CRC64:

Query Match: 85.5%; Score 3116; DB 1, Length 715;

Host Local Similarity 84.1%; Pred. No. 5.8e-186;

Matches 599; Conservative 56; Mismatches 54; Indels 4; Gaps 3;

QY 2 ENCHNGVGVNVLCTI::KIPINMGVAALITSHLEDEYKKNVAFRRKPVYVLLSGELTSSQ 59

DB 6 KN:YQNGSISLSL:VVDPROPLISGVAAEMKSGSHLDEYKPNVADYRKRYVLLSGELTSSQ 64

QY 60 QVANGTAAANDSTPVKVLSEAPAVYKASSVWMSMKSTDSYVTTGATSHPTKQ 119

DB 65 QVASI:AGHIOIGVKEVLSARHPGVKASSVWMSMKSTDSYVTTGATSHPTKQ 124

QY 120 GVALQELIPELNAJISGNTSTHPLHSATRAAMIVRNILQSGYSGIFELLHATKRLANNIT 179

DB 124 GVALQELIPELNAJISGNTSTHPLHSATRAAMIVRNILQSGYSGIFELLHATKRLANNIT 184

QY 180 PLNNNIIPCLPAGCTITASGDLVPLSYLA:LLTTRPNSKAV:PT:EVLLNAERFAAGVGGFELQPKESL 239

DB 184 PLNNNIIPCLPAGCTITASGDLVPLSYLA:LLTTRPNSKAV:PT:EVLLNAERFAAGVGGFELQPKESL 244

```

QY 240 GCFPELQKREGIALVNCVAGSCMASVLPDANVIALISEVLSAIFAEVMOCKPFTDHL 299
DB 244 SGEFFELQKREGIALVNCVAGSGLASVYKLFETVNLAVLSLSTFPAEVMKGRPEFTDH 303
QY 300 THKLKHPGQIEAAAIMETIILDSGVYKAAQVHEMDPQKODRYALRTSPQWLGPOI 359
DB 304 THKLKHPGQIEAAAIMETIILDSGVYKAAQVHEMDPQKODRYALRTSPQWLGPOI 363
QY 360 EYIRSTKTEIREINSVNDNPLIDVSRKALHSGNFOGTPGIVSMNDTPLATAITGKLMF 419
DB 364 EYIRSTKTEIREINSVNDNPLIDVSRKALHSGNFOGTPGIVSMNDTPLATAITGKLMF 423
QY 420 AOFSELVNDVFNNGLPNSLSCGRNHSIDYGFKGRHIAVASYCELOPLANPVYTHVQSAE 479
DB 424 AOFSELVNDVFNNGLPNSLSTASRNPSLIDYGFKGAETIAVASYCELOPLANPVTHVQSAE 483
QY 480 QHNQVNSLGLISAKKTAEVNDILKLMSSYLVVALCOSIDLRLHEENKSTVKNITVQVA 539
DB 484 QHNQVNSLGLISAKKTAEVNDILKLMSSYLVVALCOSIDLRLHEENKSTVKNITVQVA 543
QY 540 KKVLTIMGVNGELHPSRCEKDLRYVDREYFAYIDVCSGTYPMLQKRLQVLDHALNN 599
DB 544 KKVLTIMGVNGELHPSRCEKDLRYVDREYFAYIDVCSGTYPMLQKRLQVLDHALNN 603
QY 600 GETEKNTNTSIFOKIATFEELKVLVLPKEVEGVRIAYENDTILSPNRKACRSYPIYKRV 659
DB 604 GENENKFSYVQKLEAEVEELKALPKREVSARAAVDSNSAIDNRIKRCRSYPIYKRV 663
QY 660 REELRGELTGEKVTSPGEFDFVFTAMCKGQIIDLPLECLSGNNGEPLPIC 711
DB 664 REELRGELTGEKVTSPGEFDFVFTAMCKGQIIDLPLECLSGNNGEPLPIC 715

RESULT 10
PAL1_PROVAV STANDARD: PRT: 717 AA.
AC 064963:
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
OS PAL1.
OC Prunus avium (Cherry).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=42229;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Summit;
RA Wiersma P.A., Wu Z.;
RT "A full-length cDNA for phenylalanine ammonia-lyase cloned from ripe
RT Sweet Cherry fruit (Prunus avium).";
RL (in) Plant Gene Register PCR98-184.
CC -I- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -I- CATALYTIC ACTIVITY: L-phenylalanine + trans-cinnamate + NH(3).
CC -I- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -I- PTM: CONTAINS AN ACTIVE SITE 4-METHYLDIENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annoucr/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: AF036948; AAC78457.1; -.
DR HSSP: P21310; 188F.
DR InterPro: IPR001106; Phe/His_NH3-lyase.
DR Pfam: PF00221; PAL; 1.
DR TIGRFAMs: TIGR01226; phe_am_lyase; 1.
DR ProSite: PS00488; PAL_HISTIDASE; 1.
KW lysase; phenylpropanoid metabolism; Multigene family.
FT Site 203
FT 205
FT 5-ONE (BY SIMILARITY).
SQ SEQUENCE 717 AA; 77999 MW; B8ADP90F0A0BF60B3 CRC64:

Query Match      85.5%; Score 3115; DB 1; Length 717;
Best Local Similarity 84.7%; Pred. No. 6,7e-186;
Matches 604; Conservative 50; Mismatches 53; Indels 6; Gaps 5;

QY 2 ENGHNVCNVN--ELCI-KDPLNMGVAAEALTGSHLDEYKKNVAERKRPVYKGGETLIV 58
DB 8 QNG-HKNGSVLEPLICIKKPLNMGVAAFTIKSHLDEYKPNVAEYKRPVYKLGHSLTI 66
QY 59 SAVAGIAANDSDIVVKVLESEARAGVAKSSDWMVFSMNKGTDSYGTTCGATSHHRTK 118
DB 67 SYVAAL-ATHDS-VKVELSEARAGVAKSSDWMVFSMNKGTDSYGTTCGATSHHRTK 124
QY 119 OGCALOKELIRPLNAGIFGNGTETSHLPASATRAAMIYRINTLLQYSGIRPELLEAIT 178
DB 125 QCAALQKELIRPLNAGVFGSTKESGHTLPHQATFPAALVPIINTLLQYSGIRPELLEAIT 184
QY 179 KPLNNITPCLPLRGITIASGDLVPLSTIAGLLIGRPNKSAVGPTEGVINAEKFAAAGV 238
DB 185 KPLNNITPCLPLRGITIASGDLVPLSTIAGLLIGRPNKSAVGPTEGVINAEKFAAAGV 244
QY 239 ECGFPELOPKREGIALVNCVAGSCMASVLPDANVIALISEVLSAIFAEVMOCKPFTDHL 298
DB 245 NSGFPELOPKREGIALVNCVAGSGLASVYKLFETVNLAVLSLSTFPAEVMKGRPEFTDH 304
QY 299 LTHKLKHPGQIEAAAIMETIILDSGVYKAAQVHEMDPQKODRYALRTSPQWLGPOI 358
DB 305 LTHKLKHPGQIEAAAIMETIILDSGVYKAAQVHEMDPQKODRYALRTSPQWLGPOI 364
QY 359 IEVIPSSTKTEIREINSVNDNPLIDVSRKALHSGNFOGTPGIVSMNDTPLATAITGKLM 418
DB 365 IEVIPSSTKTEIREINSVNDNPLIDVSRKALHSGNFOGTPGIVSMNDTPLATAITGKLM 424
QY 419 FAOFSELVNDVFNNGLPNSLSCGRNHSIDYGFKGRHIAVASYCELOPLANPVYTHVQSAE 478
DB 425 FAOFSELVNDVFNNGLPNSLSCGRNHSIDYGFKGAETIAVASYCELOPLANPVYTHVQSAE 484
QY 479 FOHNDVNSLGLISAKKTAEVNDILKLMSSYLVVALCOSIDLRLHEENKSTVKNITVQVA 538
DB 485 FOHNDVNSLGLISAKKTAEVNDILKLMSSYLVVALCOSIDLRLHEENKSTVKNITVQVA 544
QY 539 AKVLTIMGVNGELHPSRCEKDLRYVDREYFAYIDVCSGTYPMLQKRLQVLDHALNN 598
DB 545 AKVLTIMGVNGELHPSRCEKDLRYVDREYFAYIDVCSGTYPMLQKRLQVLDHALNN 604
QY 599 NGETEKNTNTSIFOKIATFEELKVLVLPKEVEGVRIAYENDTILSPNRKACRSYPIYKRV 658
DB 605 NGETEKNTNTSIFOKIATFEELKVLVLPKEVEGVRIAYENDTILSPNRKACRSYPIYKRV 664
QY 659 VEEELRGELTGEKVTSPGEFDFVFTAMCKGQIIDLPLECLSGNNGEPLPIC 711
DB 665 VEEELRGELTGEKVTSPGEFDFVFTAMCKGQIIDLPLECLSGNNGEPLPIC 717

RESULT 11
PAL2_TOHAC STANDARD: PRT: 712 AA.
AC P35513; Q22114;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
GN PALA.

```

05 Nicotiana glauca (Common tobacco)
 06 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 07 Spermatophyta: Magnoliophyta, eudicotyledons: core eudicots:
 08 Asterales, easterids I, Solanales, Solanaceae, Nicotiana.
 09 NCHL_taxid=4097;
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

```

CC EMBL: M90692: AAA34176.1;
DR HSSP: P21310: 1B8F.
DR InterPro: IPR001106, Phe/His_NH3lyase.
DR Pfam: PF00221: PAL, 1.
DR TIGRfams: TIGR01226: phe_am_lyase, 1.
DR Prosite: PS00484: PAL_HISTIDASE, 1.
KW Lyase: Phenylpropanoid metabolism: Multigene family.
FT SITE 207 209 MODIFIED TO FORM 4-METHYLDIENE-IMIDAZOLE-5-ONE (BY SIMILARITY).
FT SEQUENCE 721 AA: 78495 MW: 3184772693015087 CRC64;

Query Match 85.0%; Score 3097.5; DB 1; Length 721;
Best Local Similarity 83.4%; Pred. No. 8.3e-185;
Matches 598; Conservative 50; Mismatches 62; Indels 7; Gaps 3;

OY 1 MNGNHVNGVYNELC-----IKDPLNMGVAEALTGSHLDVKKMAEPKPVKKGGET 55
DB 6 VONG-HVNGEAMDLCKKSTINNDPLNMEAAESLRGSHLDVKKMAEPKPVKKGGET 64
OY 56 LTVSQVAGIA-AANDSDTVKVELSEARAGYKASSDWMESEMNKGTDSYGVTCGATSH 114
DB 65 LTVAGVASIANVDNKNKNGVVELSPARAGYKASSDWMDCKGCTDSYGVTCGATSH 124
OY 115 RRTKGAGLQKELIFLNLNAGFNGTETSHLPHSATRAAMIVINLLQGYGSIREFIL 174
DB 125 RRTKNGAGLQKHLIKPLNAGVFGNCTESSHTLPHSATRAAMLVRIINLLQGYGSIREFIL 184
OY 175 EATKFLNNNTIPCLPLKGTITASGDVPLSYAGILTGFRNSKAVGPTGCVLAKEAFA 234
DB 185 EATKFLNNNTIPCLPLKGTITASGDVPLSYAGILTGFRNSKAVGPTGCVLAKEAFA 244
OY 235 AACGEGPPIIOPKCLALVNGTAVGSMASVLPANVALLSEVLSAIPAEVMDCKPE 294
DB 245 VAGVTSGEFELQPKGIALVNGTAVGSMASVLPFESNILLAVMESEVLSAIPAEVMDCKPE 304
OY 295 FTDHLLHKHKHPGQIEAAALMEYLLDGSQVYKAAOKVHEMDPLQPKQDQYALFRTSPQW 354
DB 305 FTDYLLHKHKHPGQIEAAALMEYLLDGSQVYKAAOKVHEMDPLQPKQDQYALFRTSPQW 364
OY 355 LGDQIEVIRSSTKMIREKINSVNDPLIDVSRNKAHCGNPGTPIGVSMNTRIALAA1 414
DB 365 LGDQIEVIRSSTKMIREKINSVNDPLIDVSRNKAHCGNPGTPIGVSMNTRIALAA1 424
OY 415 GLTMAQPSSELYNDVYNNGLPSNLSCGRNPISDYCFKQGLIAMAASYSELIQPLANPTNH 474
DB 425 GLTMAQPSSELYNDVYNNGLPSNLSCGRNPISDYCFKQGLIAMAASYSELIQPLANPTNH 484
OY 475 VQSAEQHNDVNSGLISARKTAEAVDILKLMSSSTYVALACOSIDLPHLEENMKSTYKNT 534
DB 485 VQSAEQHNDVNSGLISARKTAEAVDILKLMSSSTYVALACOSIDLPHLEENMKSTYKNT 544
OY 535 VQVAKKAVLTMGVNGELHPSRPECEDILRVYDREYVAYIDVCGSTGYPLMOKILOVLD 594
DB 545 VQVAKKAVLTMGVNGELHPSRPECEDILRVYDREYVAYIDVCGSTGYPLMOKILOVLD 604
OY 595 HALNGETEKNTNSTIPKIAFFEEELKVLPLKVPVGVRIAYENDTISIPNRIKACSYSP 654
DB 605 HALNGETEKNTNSTIPKIAFFEEELKVLPLKVPVGVRIAYENDTISIPNRIKACSYSP 664
OY 655 LYRFVREELGSGRLTGERVTSPEEFEDRTAMCKGQIIDPLELCELGWNGEPLTIC 711
DB 665 LYRFVREELGSGRLTGERVTSPEEFEDRTAMCKGQIIDPLELCELGWNGEPLTIC 721

RESULT 13
PALY_DIGLA STANDARD: PRT; 713 AA.
AC 023924:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).

```

```

OS Digitalis lanata (Foxglove).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
OC Assteridac, euasterids I, Lamiales, Veroniceae, Digitalis.
OX NCBI_TaxId=49450;
RN [1]
RP SEQUENCE FROM N.A.
RA Theoringer C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLALANINE SKELETON
CC - CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
CC - PATHWAY: Phenylpropanoid biosynthesis; first step.
CC - SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC - PTM: CONTAINS AN ACTIVE SITE 4-METHYLDIENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ002221: CAA05251.1;
DR HSSP: P21310: 1B8F.
DR InterPro: IPR001106, Phe/His_NH3lyase.
DR Pfam: PF00221: PAL, 1.
DR TIGRfams: TIGR01226: phe_am_lyase, 1.
DR Prosite: PS00484: PAL_HISTIDASE, 1.
KW Lyase: Phenylpropanoid metabolism: Multigene family.
FT SITE 198 200 MODIFIED TO FORM 4-METHYLDIENE-IMIDAZOLE-5-ONE (BY SIMILARITY).
FT SEQUENCE 713 AA: 77732 MW: 747064655DDC66A CRC64;

Query Match 85.0%; Score 3096; DB 1; Length 713;
Best Local Similarity 84.0%; Pred. No. 1e-184;
Matches 600; Conservative 50; Mismatches 56; Indels 8; Gaps 4;

OY 1 MNGNHVNGVYNELCIR--DPLNMGVAEALTGSHLDVKKMAEPKPVKKGGETLV 58
DB 5 VENGHGN---NGPCVKNDDPLNMVAAAEELKSHLDVKKMAEPKPVKKGGETLV 61
OY 59 VQVAGIAANDSDTVKVELSEARAGYKASSDWMESEMNKGTDSYGVTCGATSHRPTK 118
DB 62 SVVAALAAKDE--VAVGLAESSKAVGKASSDWMESEMNKGTDSYGVTCGATSHRPTK 119
OY 119 OGCAIQKELIFLNLNAGFNGTETSHLPHSATRAAMIVRIINLLQGYGSIREFIL 178
DB 120 OGCAIQKELIFLNLNAGFNGTETSHLPHSATRAAMLVRIINLLQGYGSIREFIL 179
OY 179 KFLNNITPCPLKGTITASGDVPLSYAGILTGFRNSKAVGPTGCVLAKEAFAAAG 238
DB 180 KFLNNITPCPLKGTITASGDVPLSYAGILTGFRNSKAVGPTGCVLAKEAFAAAG 239
OY 239 EGG-FFELQPREGLALVNGTAVGSMASVLPANVALLSEVLSAIPAEVMDCKPEFTD 297
DB 240 NSGLEFFELQPREGLALVNGTAVGSMASVLPANVALLSEVLSAIPAEVMDCKPEFTD 299
OY 298 HLTMLKHKHPGQIEAAALMEYLLDGSQVYKAAOKVHEMDPLQPKQDQYALFRTSPQW 357
DB 300 HLTMLKHKHPGQIEAAALMEYLLDGSQVYKAAOKVHEMDPLQPKQDQYALFRTSPQW 359
OY 358 QIEVIRSSTKMIREKINSVNDPLIDVSRNKAHCGNPGTPIGVSMNTRIALAAIGKI 417
DB 360 QIEVIRSSTKMIREKINSVNDPLIDVSRNKAHCGNPGTPIGVSMNTRIALAAIGKI 419
OY 418 MEAFSELYNDVYNNGLPSNLSCGRNPISDYCFKQGLIAMAASYSELIQPLANPTNHVOS 477

```

```

DB 420 MFAQSHLVNDVFNNGIIPSNISQGNHPSIDYCGKSHIAMASYCSHIOPLANVITNHVS 479
QY 478 AEQNDQVNSLGLISARKRAEVDILKMSSTYVALCOSIDRLLENMSTVANTYSQ 537
DB 480 AEQNDQVNSLGLISARKRAEVDILKMSSTYVALCOSIDRLLENMSTVANTYSQ 539
QY 538 VAKKVLITMGVNGELHSRCCEKDIILVNDREYFAYIDVCGSPYIMOKIILQVILVDHAL 597
DB 540 VAKKVLITMGVNGELHSRCCEKDIILVNDREYFAYIDVCGSPYIMOKIILQVILVDHAL 599
QY 598 NGCEKNTNTSTFQKTAIFEEELKVLIPKEVEGVPIAYENDTLCSTPNPITKRSYPLYP 657
DB 600 KNGHNKKNASTSTFQKTAIFEEELKVLIPKEVEGVPIAYENDTLCSTPNPITKRSYPLYP 659
QY 658 FVBEELARGLTGCKVTSPEEFDERETAMCGOITIDELLETLCGNNDEPPLIC 711
DB 660 FVBEELARGLTGCKVTSPEEFDERETAMCGOITIDELLETLCGNNDEPPLIC 713

RESULT 14
PALY_STYHU
ID PALY_STYHU STANDARD: PRT: 715 AA
AC 845732:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
GN PALY1.1.
OS Stylosanthes humilis (Townsille stylo).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucrotsid 1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Stylosanthes.
OX NCBI_TaxID=35628;
RN [1]
RP SOURCE: FROM N.A.
RC STRAIN: cv. Paterson; TISSUE: Stem;
RA MEDLINE=95357421; PubMed=7610950;
RX Manners J.M., McIntyre C.L., Nourse J.P.;
RT "Cloning and sequence of a cDNA encoding phenylalanine ammonia-lyase
RT from the tropical forage legume Stylosanthes humilis.";
RI Plant Physiol. 108:1301-1302(1995).
CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -1- CATALYTIC ACTIVITY: L-phenylalanine - trans-cinnamate + NH(3).
CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable)
CC -1- PTM: CONTAINS AN ACTIVE SITE: 4-METHYLENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-ch).
CC
DB HML: 136822: AAA9500.1:
DR HMLP: P21310: 1B8F.
DR InterPro: IPR001105: Phe/His_NH3Lase.
DR Pfam: PF00221: PAL_1.
DR TIGRPFAMs: TIGR01226: Phe.am.lyase.1.
DR PROSITE: PS00488: PAL_HISTIDASE_1.
KM lyase; phenylpropanoid metabolism;
FT SITE 201 203 MODIFIED TO FORM 4-METHYLENE-IMIDAZOLE-
FT SEQUENCE 715 AA: 77953 MW: 881633DE4F6A2D CEC64.
SU
Query Match 84.7%; Score 3088; D3 1; Length 715;

```

```

Host local similarity 85.7%; Pred No. 3,26-184;
Matches 595, Conservative 47, Mismatches 50, Indels 2; Gaps 2;
QY 18 EPDNLVVAEALTSIHDEFEKRVKVAEFPKPVYKLSAEFLTVSVVAIAANLSTVVEL 77
DB 24 DPLNMAAAALKSGSHIDEVKRRVSRKPIVNLGGQTLISQVAAL-ANNGQ-VSQQL 81
QY 78 SEAPAAQVAKASSVWVPSNKKGTSTGYTGGATISHPPTKGGALQKPIIPILNAGITFG 137
DB 82 SEAPAAQVAKASSVWVPSNKKGTSTGYTGGATISHPPTKGGALQKPIIPILNAGITFG 141
QY 138 NGELSHILPHSAIAAAMIVKINILUGYSGIRPELLEAIEKFINNNITPCQIRGTTA 197
DB 142 NGETENTQIPHTATRAMILVRIINTLUGYSGIRPELLEAIEKFINNNITPCQIRGTTA 201
QY 198 SIDLVLPLSYASLITLTPNSKAVPTPEVLAENKAFVAAVSEVFEELDPEGLAVNGT 257
DB 202 SGDIPLVPLSYAGLITGPRPSKAVGPNGHITNAKFAVQAGISNIFPEIOPREGLAVNGT 261
QY 258 AVGSCMASVILPIANVIALISSEVLSAIFAEVMOGRPEFDHLTHRLKHHPCGIIAAALME 317
DB 262 PVSGLSVYLFENILVAYSEVLSAIFAEVMOGRPEFDHLTHRLKHHPCGIIAAALME 321
QY 318 YIIJGSDYVKAQKVAHMDPLQPKQDRYALRTSPQWIGPQIEVIRKSTKMLREINSVN 377
DB 322 HILDGSSYVKAQKVAHMDPLQPKQDRYALRTSPQWIGPQIEVIRKSTKMLREINSVN 381
QY 378 DNLPLDVSFKKALHINFPYTPVTSVMENTPLAIALTKLMAQVSELYVNDYNNGLPSN 437
DB 382 DNLPLDVSFKKALHINFPYTPVTSVMENTPLAIALTKLMAQVSELYVNDYNNGLPSN 441
QY 438 ISCGRNLSLYGCKGEIIMASYCSHIOPLANVITNHVSARHODVNSICLISARKTA 497
DB 442 LSARKNSLDYGCKGEIIMASYCSHIOPLANVITNHVSARHODVNSICLISARKTA 501
QY 498 EAVDILKMSSTYVALCOSIDRLLENMSTVANTYSQAKKVIIMGVNCELHPSRC 557
DB 502 EAVDILKMSSTYVALCOSIDRLLENMSTVANTYSQAKKVIIMGVNCELHPSRC 561
QY 558 EKDLIPVYVPEVEYFAYIDVCGSPYIMOKIILQVILVDHALNN-ELEKNTNTSTFQKTAIF 617
DB 562 EKDLIPVYVPEVEYFAYIDVCGSPYIMOKIILQVILVDHALNN-ELEKNTNTSTFQKTAIF 621
QY 618 FEEELKVLIPKEVEGVPIAYENDTLCSTPNPITKRSYPLYPKEVEGVPIAYENDTLCSTPNP 677
DB 622 FEEELKVLIPKEVEGVPIAYENDTLCSTPNPITKRSYPLYPKEVEGVPIAYENDTLCSTPNP 681
QY 678 FEEEDRVETAMCGOITIDELLETLCGNNDEPPLIC 711
DB 682 FEEEDRVETAMCGOITIDELLETLCGNNDEPPLIC 715

RESULT 15
PAL1_LITTER
ID PAL1_LITTER STANDARD: PRT: 710 AA.
AC 049835:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5) (PAL-1).
OS Lithospermum erythrorhizon.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; euasterids I; Invertidae sedis; Boraginaceae;
OC Lithospermum.
OX NCBI_TaxID=34254;
RN [1]
RP SOURCE: FROM N.A.
RC STRAIN: 98101947; PubMed=9418980;
RX Yasaki K., Katsuka M., Honda G., Severin K., Heide L.;
RA "cDNA cloning and gene expression of phenylalanine ammonia-lyase in
RT Lithospermum erythrorhizon.";
RI Biosci. Biotechnol. Biochem. 61:1995-2003(1997).

```

```

CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -1- CATALYTIC ACTIVITY: L-phenylalanine -> trans-cinnamate + NH(3).
CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN ROOTS.
CC -1- PFM: CONTAINS AN ACTIVE SITE 4 METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announc/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D83075; BAA24928.1; -.
CC DR HSBP; P21310; 188F.
CC DR InterPro: IPR001106; Phe/HIS_NH3lyase.
CC DR pfam: PF00221; PAL; 1.
CC DR TIGRfams: TIGR01226; phe_am_lyase; 1.
CC DR PROSITE: PS00488; PAL_HISTIDASE; 1.
CC KW Lyase; Phenylpropanoid metabolism; Multigene family.
CC FT SITE 196 198 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-
CC 5-ONE (BY SIMILARITY).
CC SO SQUONCH: 710 AA; 77518 MW; CEF1D64B3A80D03 CRC64;

Query Match 84.5%; Score 3077.5; DB 1; Length 710;
Best Local Similarity 83.5%; Pred. No. 1.4e-183;
Matches 594; Conservative 53; Mismatches 59; Indels 5; Gaps 2;

QY 1 MNGNHNVCNVNVEICINPIMNGVAEALITGSHIDEVKKVAEPRKPYVKIGCTETLVSO 60
DB 5 VENG---NCKTMEFCMKDDLNEMASSESMSKSHIDEVKNMVAEPRKPYVOLACTLTIGQ 61
QY 61 VAGIAAANDSDIVKVELSEAAKAGYKASSDWVMESMKNKGTDSYGVTTGFGATSHRTKQG 120
DB 62 VAAIARDDG--VIVELAAAREGVKASSDWVMSMKNKGTDSYGVTTGFGATSHRTKQG 119
QY 121 GALQKELLIRFLNAGTFCNGTETSHTLPHSATRAAMIVRINTLDOYSGIRFELIATTK 180
DB 120 GALQKELLIRFLNAGTFCNGTETSHTLPHSATRAAMIVRINTLDOYSGIRFELIATTK 179
QY 181 LNNNITPCILPGCTTASGDVPLSYIAGLTGCRPNKAVGPTGEVLNAEKAPAAAGVPG 240
DB 180 LNNITPCILPGCTTASGDVPLSYIAGLTGCRPNKAVGPTGEVLNAEKAPAAAGVPG 239
QY 241 GFEEIQPKRGALVNGTAVGSGMASMVLFDANVLLALSEVLSAIFAEVMQKPEETDHLT 300
DB 240 GFEEIQPKRGALVNGTAVGSGMASMVLFDANVLLALSEVLSAIFAEVMQKPEETDHLT 299
QY 301 HKLKHHPGQIFAAALMEYLIDGSDYVKAOKVHENDPLOKPKODRYALRTSPQWLGPQIE 360
DB 300 HKLKHHPGQIEAAALMEYLIDGSDYVKAOKVHENDPLOKPKODRYALRTSPQWLGPQIE 359
QY 361 VIRSSTKMLIRFINSVNDPLIDVSRNKALHCGNFQGTICVSMIDNIRLAIAIGKLMPA 420
DB 360 VIRSATKMLIRFINSVNDPLIDVSRNKALHCGNFQGTICVSMIDNIRLAIAIGKLMPA 419
QY 421 QFSELYNDFYNNGLPSNLSSGGRNPSLDYGFKGGEIAAASYCELOFLANPYTNIVQSAEQ 480
DB 420 QFSELYNDFYNNGLPSNLSSGGRNPSLDYGFKGGEIAAASYCELOFLANPYTNIVQSAEQ 479
QY 481 HHQDVNSGLTISARKTAFAVDLKLMSSTYIALCOSIDLRLPENKSTYKNTVSQVAK 540
DB 480 HHQDVNSGLTISARKTAFAVDLKLMSSTYIALCOSIDLRLPENKSTYKNTVSQVAK 539
QY 541 KVLTMGVNGELHPSRFCKKDLLRVVDREYPAVIDVCSCTYPLMOKLROVLYDHAIANG 600
DB 541 KVLTMGVNGELHPSRFCKKDLLRVVDREYPAVIDVCSCTYPLMOKLROVLYDHAIANG 600

```

```

DB 540 RLTTGVNGELHPSRFCKKDLLRVVDREYPAVIDVCSCTYPLMOKLRETLVGHAIANG 599
QY 601 EREKNTNNSIFQKITEFEELKVLIPKEVEGVPIAYENDLTSPNPIKARSTYPLTFVA 660
DB 600 ENEKDVNTSIFKIAIFEEELKAILPEVENARASVENGIPALISNRLECRSTYPLTFVA 659
QY 661 EELGKFTLGEKVTSPGEFPRVFTAMCKQIIPILLECICGWNCEPIPIIC 711
DB 660 EELGTELLTGEKVTSPGEELDKVFTAMCESKLVDPPLACLEANNGAPLPIIC 710

```

Search completed: March 29, 2003, 02:10:56
Job time: 42 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mod-1

Run on: March 29, 2003, 01:54:51, Search time 63 Seconds

(without alignments)
2325.389 Million cell updates/sec

Title: US-09-964-992A-1
Perfect score 3644
Sequence: 1 MENCNHNVCVNEELTKDPL 110PLPVTATWNEPLPTG 711

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

1: SPREMBL_21:*
2: sp-archaea:*
3: sp-bacteria:*
4: sp-fungi:*
5: sp-human:*
6: sp-invertebrate:*
7: sp-mammal:*
8: sp-organellar:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriophage:*
17: sp-archaeo:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3644	100.0	711	10	Q8W2E4
2	3195.5	87.7	716	10	Q9MAX1
3	3148.5	86.4	730	10	Q9M567
4	3107.5	85.3	708	10	Q94EN0
5	3106	85.2	721	10	Q9XF55
6	3100.5	85.1	711	10	Q9AX15
7	3093	84.9	710	10	Q94C45
8	3090	84.8	716	10	Q94AT7
9	3084.5	84.6	712	10	Q94F89
10	3056.5	83.9	718	10	Q94FX6
11	3049	83.7	715	10	Q94J66
12	3045.5	83.6	725	10	Q94A41
13	3009	82.6	717	10	Q8RW44
14	2995	82.2	687	10	Q94F90
15	2820.5	77.4	655	10	Q9XG83
16	2671	73.3	619	10	Q95455

17	2645	72.6	618	10	Q9FS00
18	2496	64.5	703	10	Q8VX57
19	2400	65.9	532	10	Q81647
20	2332.5	64.0	681	10	Q8R023
21	2331.5	64.0	681	10	Q8R049
22	2325.5	63.8	681	10	Q8SAS7
23	2321.5	63.7	681	10	Q8SAS6
24	2190	60.1	544	10	Q9E138
25	2184	59.9	549	10	Q94876
26	1948	53.5	475	10	Q94875
27	1887	51.8	497	10	Q94869
28	1855	50.9	496	10	Q8WVP6
29	1737	47.7	396	10	Q9SWP6
30	1693	46.5	362	10	Q9F017
31	1675	46.0	363	10	Q9F018
32	1616	44.3	356	10	Q9S0M2
33	1573	43.2	395	10	Q8S3W3
34	1368	37.5	339	10	Q43211
35	1365	37.5	315	10	Q94F91
36	1355	37.2	319	10	Q80406
37	1292.5	35.5	436	10	Q91E71
38	1288.5	35.4	435	10	Q9M301
39	1179	32.4	285	10	Q91E82
40	1063	29.2	240	10	Q940D8
41	977	26.8	217	10	Q9M627
42	844.5	23.2	740	3	Q93967
43	833	22.9	173	10	Q82476
44	832	22.8	178	10	Q41293
45	811	22.3	186	10	Q9M626

ALIGNMENTS

RESULT 1	Q8W2E4	PRELIMINARY:	PRT:	711 AA.
ID	Q8W2E4			
AC	Q8W2E4:			
DT	01-MAR-2002 (STEMBL_21, created)			
DT	01-MAR-2002 (STEMBL_21, last sequence update)			
DT	01-JUN-2002 (STEMBL_21, last annotation update)			
DE	Phenylalanine ammonia-lyase.			
OS	Lactuca sativa (garden lettuce).			
OC	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta:			
OC	Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots:			
OC	Asteridae, euasterids II, Asterales, Asterales, Lactuca:			
OX	NCBI_TaxID=4236;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
PA	Chompos-Vargas R., Nonogaki H., Suslow T., Salviati M.:			
FT	"Characterization of phenylalanine ammonia-lyase (PAL) gene in wounded lettuce leaf tissue."			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF299330; AML55242.1;			
DR	Uniprot: IPR001106; Phe/His_NH3lyase.			
DR	Plan: PF00221; PAL, 1.			
DR	PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN_1.			
KW	LYASE.			
SC	SEQUENCE 711 AA: 77363 MW: 3120A82877C29513 CMC64:			
Query Match	100.0%; Score 3644; DB 10; Length 711;			
First Local Similarity	100.0%; Pred Nc 9.7e-233;			
Matches 711, Conservative 0, Mismatches 0, Indels 0, Gaps 0;				
QY	1 MENCNHNVCVNEELTKDPLNNVAAAEALTGSHLDVKKKVAEPKPVKLGCTTVSQ 60			Q91S06 dianthus ca
Pb	1 MENCNHNVCVNEELTKDPLNNVAAAEALTGSHLDVKKKVAEPKPVKLGCTTVSQ 60			Q8VX57 cea mayas (m
QY	61 VACIAAANDSDIVKVLSPAAAVGKASSDWMPSMNGGTHSVGVTGCGATSRPPIKQG 120			Q81647 copistum ch
Ub	61 VACIAAANDSDIVKVLSPAAAVGKASSDWMPSMNGGTHSVGVTGCGATSRPPIKQG 120			Q8R023 pinus sylve
QY	121 GALKELIRLNAGTGTGNTSTSHLPISATATAAAVHNTLLGGYSIGIRFELLEATIKF 180			Q8R049 pinus sylve

```

|||||
DB 121 GAGGKELIIRFLNACIFGCTEFSHTLPHSATRAAMIVINILLOCGYSIGREILLEATITK 180
QY 181 LNNNTTPELFLEGTASGLVPLSTAGLLTIRKNSKAV:PTSEVINAERKAFAAVEG 240
DB 181 LNNNTTPELFLEGTASGLVPLSTAGLLTIRKNSKAV:PTSEVINAERKAFAAVEG 240
QY 241 GFEELQPKKGLAVNGTAVGSGMASMVLFDANVIALLESEVISAIFAEVYMGKPEFTDHLT 300
DB 241 GFEELQPKKGLAVNGTAVGSGMASMVLFDANVIALLESEVISAIFAEVYMGKPEFTDHLT 300
QY 301 HKIKHHPCQIEAAAIMEVYILDCSDYVKAOKVHEMDPLQKPKQDRAVALRTSPQW:GPOIE 360
DB 301 HKIKHHPCQIEAAAIMEVYILDCSDYVKAOKVHEMDPLQKPKQDRAVALRTSPQW:GPOIE 360
QY 361 VRSSTKMTIEREINSVNDNPLIDVSRNKALHGCNFOGTPICVSMNDNRLAIAAGKLMFA 420
DB 361 VRSSTKMTIEREINSVNDNPLIDVSRNKALHGCNFOGTPICVSMNDNRLAIAAGKLMFA 420
QY 421 QFSELVNFYNNGLPSNLISGGKNPSLDYGFKGCEIAMAASYCELOFLANPVTNHVQSAEQ 480
DB 421 QFSELVNFYNNGLPSNLISGGKNPSLDYGFKGCEIAMAASYCELOFLANPVTNHVQSAEQ 480
QY 481 HNOOVNSLGLISAKTAFAVDILKLMSSYLVALCOSTIDLRHLENNKSTVKNVSOVAK 540
DB 481 HNOOVNSLGLISAKTAFAVDILKLMSSYLVALCOSTIDLRHLENNKSTVKNVSOVAK 540
QY 541 KVLTVGNGELHPSRCEKDLIRVVDREYFAYIDVCSCTPYLMOKIROVYDHALNNG 600
DB 541 KVLTVGNGELHPSRCEKDLIRVVDREYFAYIDVCSCTPYLMOKIROVYDHALNNG 600
QY 601 ETEKNTNISFOKIAATFEELKVLTPKEVEGRIAYENDTISINRIKACSYPLRYFVR 660
DB 601 ETEKNTNISFOKIAATFEELKVLTPKEVEGRIAYENDTISINRIKACSYPLRYFVR 660
QY 661 EELGKPLTGEKVTSPGCEFDVFTAMCKGQIIDPLLECLCGMNGEPLPIC 711
DB 661 EELGKPLTGEKVTSPGCEFDVFTAMCKGQIIDPLLECLCGMNGEPLPIC 711

RESULT 2
ID Q9MAX1 PRELIMINARY: PRT: 716 AA.
AC Q9MAX1:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Phenylalanine ammonia-lyase.
OS Catharanthus roseus (Rosa periwinkle) (Madagascar periwinkle).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicotys:
OC Asteridae: euasterids I: Gentianales: Apocynaceae: Rauvolfioidae:
OC Vinceae: Catharanthus.
OX NCBI_TaxID=4058;
RN 11
RP SEQUENCE FROM N.A.
RA Kiyota S., Sakao K., Kim J.:
RT "cDNA of phenylalanine ammonia-lyase from Catharanthus roseus.";
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AH042520; BAA95629.1;
DR HSSP: P21310; 188F.
DR InterPro: IPR001106; Phe/His_NH3lyase.
DR Pfam: PF00221; PAL; 1.
DR TIGRfams: TIGR01226; phe_am_lyase.1.
DR PROSITE: PS00488; PAL_HISTIDASE; 1.
SQ SEQUENCE 716 AA: 78241 MW: 1F16D4DEB8683B0B CRC64;

Query Match 87.7% Score 3195.5: nr10: Length 715;
Best Local Similarity 86.9%: Pred No.2-be-194;
Matches 617: conservative 43; Mismatches 49; Indels 1; Gaps 1;
2 ENGNGVNGVNNICITKIDPIMNGVAFAEATGSHLDEYKKAWEFRKPVKICGCTTIVSQV 61

```

```

|||||
DB 8 ENGNIINNEVIDFCLKPLPLNNMGMAADSLKSHLDEYKKAWEFRKPVKICGCTTITISQV 67
QY 62 AGLIAANDSTGVKVELSENAFAVYKASSDVMVE:SNK:STLY:VTT:Q:VAT:SHRPTK:Q:3 121
DB 68 AAL-AARQYNAVKVELSHDAAGVAKASSDWMBSMNKQSDYGVTTGPGATSHRRTKQCG 126
QY 122 ALQKELIFPLNAGLFGNTESSHTLPHSATRAAMIVINILLOCGYSIGREILLEATITKFL 181
DB 122 ALQKELIFPLNAGLFGNTESSHTLPHSATRAAMIVINILLOCGYSIGREILLEATITKFL 181
QY 181 NNNITPCPLNCTTTASGDLPLSYIAGLLTGRNSKAVGPTGCEVINAKKAFMAACVEG 241
DB 181 NNNITPCPLNCTTTASGDLPLSYIAGLLTGRNSKAVGPTGCEVINAKKAFMAACVEG 241
QY 242 FFEELQPKKGLAVNGTAVGSGMASMVLFDANVIALLESEVISAIFAEVYMGKPEFTDHLT 301
DB 242 FFEELQPKKGLAVNGTAVGSGMASMVLFDANVIALLESEVISAIFAEVYMGKPEFTDHLT 301
QY 302 KIKHHPCQIEAAAIMEVYILDCSDYVKAOKVHEMDPLQKPKQDRAVALRTSPQW:GPOIEV 361
DB 302 KIKHHPCQIEAAAIMEVYILDCSDYVKAOKVHEMDPLQKPKQDRAVALRTSPQW:GPOIEV 361
QY 362 IRSSTKMTIEREINSVNDNPLIDVSRNKALHGCNFOGTPICVSMNDNRLAIAAGKLMFAQ 421
DB 362 IRSSTKMTIEREINSVNDNPLIDVSRNKALHGCNFOGTPICVSMNDNRLAIAAGKLMFAQ 421
QY 422 FSELVNDYNNGLPSNLISGGKNPSLDYGFKGCEIAMAASYCELOFLANPVTNHVQSAEQH 481
DB 422 FSELVNDYNNGLPSNLISGGKNPSLDYGFKGCEIAMAASYCELOFLANPVTNHVQSAEQH 481
QY 482 NODVNSLGLISAKTAFAVDILKLMSSYLVALCOSTIDLRHLENNKSTVKNVSOVAK 541
DB 482 NODVNSLGLISAKTAFAVDILKLMSSYLVALCOSTIDLRHLENNKSTVKNVSOVAK 541
QY 542 VLTGVNGELHPSRCEKDLIRVVDREYFAYIDVCSCTPYLMOKIROVYDHALNNG 601
DB 542 VLTGVNGELHPSRCEKDLIRVVDREYFAYIDVCSCTPYLMOKIROVYDHALNNG 601
QY 602 TEKNTNISFOKIAATFEELKVLTPKEVEGRIAYENDTISINRIKACPSYPLRYFVR 661
DB 602 TEKNTNISFOKIAATFEELKVLTPKEVEGRIAYENDTISINRIKACPSYPLRYFVR 661
QY 662 EELGKPLTGEKVTSPGCEFDVFTAMCKGQIIDPLLECLCGMNGEPLPIC 711
DB 662 EELGKPLTGEKVTSPGCEFDVFTAMCKGQIIDPLLECLCGMNGEPLPIC 711

RESULT 3
ID Q9M567 PRELIMINARY: PRT: 730 AA.
AC Q9M567:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Phenylalanine ammonia-lyase 2 (EC 4.3.1.5).
OS Rubus idaeus (Raspberry).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicotys:
OC Eurosids I: Rosales, Rosaceae, Rosoidae, Rubus.
OX NCBI_TaxID=33247;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE=21437948; PubMed=11553751;
RA Kumar A., Ellis B.E.:
RT "The phenylalanine ammonia-lyase gene family in raspberry: structure,
RT expression, and evolution.";
RL Plant Physiol. 127:230-239 (2001).
DR EMBL: AF237955; AAF0224.1;
DR HSSP: P21310; 188F.
DR InterPro: IPR001106; Phe/His_NH3lyase.
DR Pfam: PF00221; PAL; 1.

```

DR TIGR01226: phe_am_lyase; 1.
DR PROSITE: PS00488, PAL_HISTIDASE, 1.
KM lysase.
SU SEQUENCE: 730 AA; 79422 MW; 98942417550CE113 CAC64;

Query Match 86.4%; Score 3148.5; DR 10; Length 730;
Best Local Similarity 83.9%; Pred. No. 2,8e-191;
Matches 610; Conservative 48; Mismatches 50; Indels 19; Gaps 4;

QY 2 ENC-NHNVGVN-----ELCIK-----DPLNMGVAEALTGSHLDVKKMAVE 44
DB 6 QNHSHHDM:QNGSLDMDGVTIKTESIKTGYSDPLNMGVAEALTGSHLDVKKMAVE 65
QY 45 RKPVNKGRETTVSQVAVIAANDSDYKVELSEAAAGVASSRWMESSNKKGDSYG 104
DB 66 KKVYKIGGHTITISQVAAI--ANHDSQKVELAFASAGVASSRWMESSNKKGDSYG 123
QY 105 VTGFSATSHRPTKQVATALQKELIRFLNATFRTNITETSHTPHATRAAMIVPINTLLQ 164
DB 124 VTGFSATSHRPTKQVATALQKELIRFLNATFRTNITETSHTPHATRAAMIVPINTLLQ 183
QY 165 GYSGIRREIILATKFLNNNTYPCUPURGTIASGDIYVLSIAGILIGRNSKAVPTG 224
DB 184 GYSGIRREIILATKFLNNNTYPCUPURGTIASGDIYVLSIAGILIGRNSKAVPTG 243
QY 225 EVLNAAKFAAAGVGVAFPELOPKREGIALVNGTAVSGVASMVLDPANVALLESELSAI 284
DB 244 FTLMAAFAAAGVGVAFPELOPKREGIALVNGTAVSGVASMVLDPANVALLESELSAI 303
QY 285 FAEVNMGKPEFTDHLTKLKHHPGQTEAAMIMEYLLDSGVYKAAQVHEMPLQPKROD 344
DB 304 FAEVNMGKPEFTDHLTKLKHHPGQTEAAMIMEYLLDSGVYKAAQVHEMPLQPKROD 363
QY 345 RVALRTSPQMLGQIIVHSSSTKMIHREINSVNDNPLDIVSRNKALHGGNFOGTIGVSM 404
DB 364 RVALRTSPQMLGQIIVHSSSTKMIHREINSVNDNPLDIVSRNKALHGGNFOGTIGVSM 423
QY 405 DNTRLIAATGKLMFAQFSELYNDYFNNGLPNSLSGGNPSLDYFGKGEIIMASVSEL 464
DB 424 DNTRLIAATGKLMFAQFSELYNDYFNNGLPNSLSGGNPSLDYFGKGEIIMASVSEL 483
QY 465 QFLANPNTNIVSAGEHNDVNSLGLISARKTAEAVIDILKMSSTYLVALCOSIDRLILE 524
DB 484 QFLANPNTNIVSAGEHNDVNSLGLISARKTAEAVIDILKMSSTYLVALCOSIDRLILE 543
QY 525 ENKSTYVKNNTVSQAKKVLTMGVNGELHPSRCCKDLRVNDREVFVATIDVCGSTPL 584
DB 544 ENKSTYVKNNTVSQAKKVLTMGVNGELHPSRCCKDLRVNDREVFVATIDVCGSTPL 603
QY 585 MOKLROVLVDHALNGETEKNTSTFQKATFEEBELKVLKPEVGVRIAYENDTSLIP 644
DB 604 MOKLROVLVDHALNGETEKNTSTFQKATFEEBELKVLKPEVGVRIAYENDTSLIP 663
QY 645 NRIKATRSYPLRYFVREBELSPFRTGEEKVTSPGSEEDRVTAMGGIIDPLELTSQWN 704
DB 664 NRIKATRSYPLRYFVREBELSPFRTGEEKVTSPGSEEDRVTAMGGIIDPLELTSQWN 723
QY 705 GEPPLPIC 711
DB 724 GEPPLPIC 730

RESULT 4
Q94FNO PRELIMINARY: PRT: 708 AA.

AC Q94FNO: 01-DEC-2001 (TIGR01226, 19, Created)
DT 01-DEC-2001 (TIGR01226, 19, Last sequence update)
DT 01-JUN-2002 (TIGR01226, 21, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
GN PAL1.
OS Rhamnia glutinosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Streptophyta; Embryophyta; Eudicotyledons; Core eudicots;
OT Asteridae; Euasterids I; Lamiales; Lamiales; Invertase; Rhamnia;
UX NCHI_taxid=99300;

RP SEQUENCE FROM N.A.
RA STRAIN:CV KEMISAN.

RA Yun S.J., Lee B.K., Park M.K., Shin J.B., Choo J.-C.;
RT "Molecular characterization of phenylalanine ammonia-lyase gene in
Rhamnia glutinosa";

RI Submitted (JUN-2001) to the EMBL/Genbank/DDBJ databases.
IR EMBL: AF401636; AAK84225.1;

DR InterPro: IPR001105; Phe/His_NH3lyase.
DR InterPro: IPR003880; Prantase_attach.

DR Pfam: PF00221; PAL_1.
DR TIGR01226: phe_am_lyase; 1.

DR PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN; 1.
DR Lysase; PS00012; PHOSPHOPANTETHEINASE; UNKNOWN; 1.

SU SEQUENCE 708 AA; 76835 MW; 186AF0816E1BAFE CAC64;

Query Match 85.3%; Score 3107.5; DR 10; Length 708;
Best Local Similarity 84.3%; Pred. No. 1.1e-186;
Matches 602; Conservative 51; Mismatches 52; Indels 9; Gaps 4;

QY 1 MGNHNVGVNDELCT---KDLNMGVAEALTGSHLDVKKMAVEPKPVKLGGETLI 57
DB 1 MGNHNVGVNDELCT---KDLNMGVAEALTGSHLDVKKMAVEPKPVKLGGETLI 57
QY 58 VSOVAGIAANISDIPVKNISPAARAGVAKASSDWMESSNKKGDSYGVITGFAISHRT 117
DB 58 IAVVAAIAAFDNN--VAVELAETAFAGVASSRWMESSNKKGDSYGVITGFAISHRT 115
QY 118 KOGGALQKELIRFLNATFONGTTSHTLPHSATRAAMIVPINTLLQYSGIRREIILAI 177
DB 116 KOGGALQKELIRFLNATFONGTTSHTLPHSATRAAMIVPINTLLQYSGIRREIILAI 175
QY 178 TKELNNNTYPCUPURGTIASGDIYVLSIAGILIGRNSKAVPTGSEVLENAEFAAAG 237
DB 176 TKELNNNTYPCUPURGTIASGDIYVLSIAGILIGRNSKAVPTGSEVLENAEFAAAG 235
QY 238 VEGGFPELOPKREGIALVNGTAVSGVASMVLDPANVALLESELSAIFAFVYKQKPEFTD 297
DB 236 V-EGGFPELOPKREGIALVNGTAVSGVASMVLDPANVALLESELSAIFAFVYKQKPEFTD 294
QY 298 HTHKLKHHPGQTEAAMIMEYLLDSGVYKAAQVHEMPLQPKRODYALRTSPQMLGP 357
DB 295 HTHKLKHHPGQTEAAMIMEYLLDSGVYKAAQVHEMPLQPKRODYALRTSPQMLGP 354
QY 358 QIEVIRPSTKMIHREINSVNDNPLDIVSRNKALHGGNFOGTIGVSMONTPLIAATGKL 417
DB 355 QIEVIRPSTKMIHREINSVNDNPLDIVSRNKALHGGNFOGTIGVSMONTPLIAATGKL 414
QY 418 MPAQFSELYNDYFNNGLPNSLSGGNPSLDYFGKGEIIMASVSELQFLANPNTNIVS 477
DB 415 MPAQFSELYNDYFNNGLPNSLSGGNPSLDYFGKGEIIMASVSELQFLANPNTNIVS 474
QY 478 APOHNDVNSLGLISARKTAEAVIDILKMSSTYLVALCOSIDRLILEHLENNKSTVNTVSQ 537
DB 475 APOHNDVNSLGLISARKTAEAVIDILKMSSTYLVALCOSIDRLILEHLENNKSTVNTVSQ 534
QY 538 VAKKVLTMGVNGELHPSRCCKDLRVNDREVFVATIDVCGSTPLMOKLROVLVDHAL 597
DB 535 VAKKVLTMGVNGELHPSRCCKDLRVNDREVFVATIDVCGSTPLMOKLROVLVDHAL 594
QY 598 NNGETEKNTSTFQKATFEEBELKVLKPEVGVRIAYENDTSLIPNIAACRSYPLRYK 657
DB 595 NNGETEKNTSTFQKATFEEBELKVLKPEVGVRIAYENDTSLIPNIAACRSYPLRYK 654
QY 658 FVREBELKVLKPEVGVRIAYENDTSLIPNIAACRSYPLRYK 711
DB 655 FVREBELKVLKPEVGVRIAYENDTSLIPNIAACRSYPLRYK 708


```

QY 480 QHNDVNSLILISAKTIAEVDILKMSSTYLVALVUSIDLRHLEENKMSSTYKNTVSQVA 539
DB 480 QHNDVNSLILISAKTIAEVDILKMSSTYLVALVUSIDLRHLEENKMSSTYKNTVSQVA 539
QY 540 KKVLTMSVNGELIPSECEKDLRVDPREYFAYIDNVCSGYPLMOKLROVLYDHALNN 599
DB 540 KKVLTMSVNGELIPSECEKDLRVDPREYFAYIDNVCSGYPLMOKLROVLYDHALNN 599
QY 600 GEFTEKNTNISFOKIAFPEHEIKVILIPKRVBCVRIAYENDLISINRIKACRSYLYRNV 659
DB 600 GESEKNTGSTIFLKVAFADEDEKAVLPKEVFAPIAVESGNPAITNPRIKEVPSYPLKRV 659
QY 660 FEELSKFSTGEKVTSPGEELRVFTAMCKSLIDPLELLELGGMSPEPLPIC 711
DB 660 KRGICITRILITGEKRVSGHPCIKVPIAMCGOSIIDLILKLSKMDCAPIPIIC 711

RESULT 7
Q94C45 PRELIMINARY: PRI: 710 AA.
AC Q94C45
DT 01-DEC-2001 (TREMblrel, 19, Created)
DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
CN PAL1.
OS Manihot esculenta (Cassava) (Manihot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Rosidae;
OC Eurosid 1; Malpighiales; Euphorbiaceae; Manihot.
OX NCBI_TaxID=3983;
RN 111
RP SOURCE FROM N.A.
RA Han Y., Boeching J.R.;
RT "Cassava phenylalanine ammonia-lyase expressed during post-harvest
RT physiological deterioration".
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY036011; AAK62030.1;
DR InterPro: IPR001106; Phc/Hls_NH3lyase.
DR Pfam: PF00221; PAL; 1.
DR TIGRPFAM: TIGR01226; Phc-am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN; 1.
KW Lyase.
SQ SHOUNCE 710 AA: 77559 MW: F236AEH1PZHA1CSC CRC64.

Query Match 84.98; Score 3093; DR 10; Length 710;
Best Local Similarity 85.38; Pred. No. 9e-188;
Matches 592; Conservative 52; Mismatches 48; Indels 2; Gaps 1.
QY 18 DPLNMCVAAHAIQSHLDHYKKKVAEPKPVYKLGFEITVVSQVAGIAAANDSDTVKVI. 77
DB 18 DPLNMCVAAHAIQSHLDHYKKKVAEPKPVYKLGFEITVVSQVAGIAAANDSDTVKVI. 77
QY 19 DPLNMCVAAHAIQSHLDHYKKKVAEPKPVYKLGFEITVVSQVAGIAAANDSDTVKVI. 76
DB 19 DPLNMCVAAHAIQSHLDHYKKKVAEPKPVYKLGFEITVVSQVAGIAAANDSDTVKVI. 76
QY 78 SEARARAVKAKSSQWVMEKMKGTSTGYVTTGPAITSHPKQVSLDQELLIFLNAITPG. 137
DB 78 SEARARAVKAKSSQWVMEKMKGTSTGYVTTGPAITSHPKQVSLDQELLIFLNAITPG. 137
QY 77 SEARARAVKAKSSQWVMEKMKGTSTGYVTTGPAITSHPKQVSLDQELLIFLNAITPG. 136
DB 77 SEARARAVKAKSSQWVMEKMKGTSTGYVTTGPAITSHPKQVSLDQELLIFLNAITPG. 136
QY 138 NGFTSTHTPHSAFRAAMIIRINTLLQSGSIFPELLEAITFEINNVTPEPLDGLTITA. 197
DB 138 NGFTSTHTPHSAFRAAMIIRINTLLQSGSIFPELLEAITFEINNVTPEPLDGLTITA. 197
QY 137 NGFTSTHTPHSAFRAAMIIRINTLLQSGSIFPELLEAITFEINNVTPEPLDGLTITA. 196
DB 137 NGFTSTHTPHSAFRAAMIIRINTLLQSGSIFPELLEAITFEINNVTPEPLDGLTITA. 196
QY 198 SGDIWPIASYIAGLITGPNSKAVGPITGVINAKKAPAAAGVCGPPELOPKIGLAINVGT. 257
DB 198 SGDIWPIASYIAGLITGPNSKAVGPITGVINAKKAPAAAGVCGPPELOPKIGLAINVGT. 257
QY 197 SGDIWPIASYIAGLITGPNSKAVGPITGVINAKKAPAAAGVCGPPELOPKIGLAINVGT. 256
DB 197 SGDIWPIASYIAGLITGPNSKAVGPITGVINAKKAPAAAGVCGPPELOPKIGLAINVGT. 256
QY 258 AVSGSMASMTLFTANVALLSEVLSAIFAEVMSKPEFTDILTHIKKHHPQITEAATIME. 317
DB 258 AVSGSMASMTLFTANVALLSEVLSAIFAEVMSKPEFTDILTHIKKHHPQITEAATIME. 316
QY 257 AVSGGLASMTLFTANVALLSEVLSAIFAEVMSKPEFTDILTHIKKHHPQITEAATIME. 316
DB 257 AVSGGLASMTLFTANVALLSEVLSAIFAEVMSKPEFTDILTHIKKHHPQITEAATIME. 316
QY 318 YLIDGSIYVAAKQVHMDPIQPKODRYALRTSPQMLGPQLEVIKSSYKMLEIREINSVN. 377
DB 318 YLIDGSIYVAAKQVHMDPIQPKODRYALRTSPQMLGPQLEVIKSSYKMLEIREINSVN. 377
QY 317 HVLDRSSYIAAOKVHFIPPLQPKODRYALRTSPQMLGPQLEVIKSSYKMLEIREINSVN. 376
DB 317 HVLDRSSYIAAOKVHFIPPLQPKODRYALRTSPQMLGPQLEVIKSSYKMLEIREINSVN. 376

```

```

QY 578 DNLIDVSRKALHGNFGITPLGVSMNTKLALAIAGLKHAFSELYNDFTYNNGLPSN. 437
DB 377 DNLIDVSRKALHGNFGITPLGVSMNTKLALAIAGLKHAFSELYNDFTYNNGLPSN. 436
QY 438 LSSAPNSCLSYSGKSEIEMASVCSFLCANPVIINWVSGEJHGVNSVSLTAPPTA. 497
DB 437 LSSAPNSCLSYSGKSEIEMASVCSFLCANPVIINWVSGEJHGVNSVSLTAPPTA. 496
QY 498 EAVDILKMSSTYLVALVUSIDLRHLEENKMSSTYKNTVSQVAKVILMGVNGELIPSEFC. 557
DB 497 EAVDILKMSSTYLVALVUSIDLRHLEENKMSSTYKNTVSQVAKVILMGVNGELIPSEFC. 556
QY 558 EKDLKVVDPREYFAYIDNVCSGYPLMOKLROVLYDHALNNGEERTKNTNISFOKIAF. 617
DB 557 EKDLKVVDPREYFAYIDNVCSGYPLMOKLROVLYDHALNNGEERTKNTNISFOKIAF. 616
QY 618 EELKYLPEVEGVPIAVENDLISINRIKACRSYLYRNVFEELGGEFTGEKVTSPG. 677
DB 617 EELKYLPEVEGVPIAVENDLISINRIKACRSYLYRNVFEELGGEFTGEKVTSPG. 676
QY 678 EEDRVFTAMCKGQIIDLPLELLELGGMSPEPLPIC 711
DB 677 EEDRVFTAMCKGQIIDLPLELLELGGMSPEPLPIC 710

RESULT 8
Q94N17 PRELIMINARY: PRI: 710 AA.
AC Q94N17
DT 01-JUN-2001 (TREMblrel, 17, Created)
DT 01-JUN-2001 (TREMblrel, 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Phenylalanine ammonia-lyase.
CN PAL1.
OS Agastache rugosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
OC Asteroideae; easteroids 1; Lamiales; Lamiales; Agastache.
OX NCBI_TaxID=39271;
RN 111
RP SOURCE FROM N.A.
RA Kim J.-B., Kim J.-B., Cho K.-J., Moon U., Ha S.-H.;
RT "Molecular cloning and characterization of Phenylalanine Ammonia-
RT lyase(PAL) Related to the Biosynthesis of Rosmarinic acid in Agastache
RT rugosa".
KL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF326116; AAK15640.1;
DR HSSF: P21310, 1B8F.
DR InterPro: IPR001106; Phc/Hls_NH3lyase.
DR Pfam: PF00221; PAL; 1.
DR TIGRPFAM: TIGR01226; Phc-am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; 1.
KW Lyase.
SQ SEQUENCE 710 AA: 77487 MW: 9403F33B8CFE2958 CRC64.

Query Match 84.88; Score 3090; DR 10; Length 710;
Best Local Similarity 83.28; Pred. No. 1.4e-187;
Matches 595; Conservative 54; Mismatches 56; Indels 8; Gaps 4;
QY 3 NGNH---VNVVNEUCIK--PPLNMVAAEALTSNLEVRKVAEPKPVYKLGFEITL. 56
DB 4 NGHGSNGNHNGFANGCPQNDPIANMAAAPSIAKQSHIFEVKVPKPVYKLGFEITL. 63
QY 57 TVSGVAGIAAANDSDTVKVIISPAARAGKASSIOWMESMKGTSTGYVTTGPAITSHP. 116
DB 64 TTSVAAIYAKTKNA--VAVELASAPAVKASSIOWMESMTSTGYVTTGPAITSHP. 121
QY 117 TKOGALOKELIFLNAITPGNCTFTSHTPHSAFRAAMIIRINTLLQSGSIFPELLEAITFE. 176
DB 122 TKOGALOKELIFLNAITPGNCTFTSHTPHSAFRAAMIIRINTLLQSGSIFPELLEAITFE. 181
QY 177 TTRFLNNNTPTLPIKTTITASGLVPLSYIANLLTGPNSKAVNPITGEVINAERKATPAAN. 236
DB 177 TTRFLNNNTPTLPIKTTITASGLVPLSYIANLLTGPNSKAVNPITGEVINAERKATPAAN. 236

```


Query Match 83.9%; Score 3056.5; DB 10; Length 718;
 Best Local Similarity 83.0%; Pred. No. 1,96,185;
 Matches 596, Conservative 47, Mismatches 60, Indels 15, Gaps 4;

QY 2 ENNRNNGVNNELTK-----DPLNMGVAEALTGSHDEVRKMAVEPRPVAKG 53
 DB 8 HNG-HQNGHGLGCKNNNNYSSGDALNMGVMAETLKSHLEVRKMAVEPRPVAKG 66

QY 54 ELLTVSVNIAAANDSVYKVELSEAAKAYKASSIWMWESNNKGTDSYVTTGKATS 113
 DB 67 ELLTVAAVAIA-----TSSNVELSESAREGVASSDWMESNNKGTDSYVTTGKATS 122

QY 114 HRTKOGAIAQKELIRFLNAGIIGNCTETSHLPHSATPAAMIVRINTLQYSGIPEEI 173
 DB 123 HRTKNGGALQKELIFLNAGIFGNCTKSHLPHSATPAAMIVRINTLQYSGIPEEI 182

QY 174 LEATKLNNTTCLPLRTTTASGDIYPLSTIAGLTGPPNSKAVPTGSEVLAKEAF 233
 DB 183 LKATKLNNTTCLPLRTTTASGDIYPLSTIAGLTGPPNSKATGPHGO 242

QY 234 AAGVEGGEFELQPKESIALVNIIVAGSGMASVLETAVALLESEVLSAIFAEVQKRP 293
 DB 243 KPAKF--GFELOPKRELALVNIIVAGSGMASVLETAVALLESEVLSAIFAEVQKRP 300

QY 294 EFTDHLTHKLRHDPQLEAAALMEVTLDSGVYKAAQVHEMDLPQKPKODRYALRTSPQ 353
 DB 301 EFTDHLTHKLRHDPQLEAAALMEVTLDSGVYKAAQVHEMDLPQKPKODRYALRTSPQ 360

QY 354 WLSPOLEVISSSTKMEIREINSVNDPLIDVSNKALHGGNFQGTPIGVSMTKLATAA 413
 DB 361 WLSPOLEVISSSTKMEIREINSVNDPLIDVSNKALHGGNFQGTPIGVSMTKLATAA 420

QY 414 IGLMEAFSELYNDEYNNGLPSNLSGGNPSLDYGFKGSEIAMASVSELOFLANPTN 473
 DB 421 IGLMEAFSELYNDEYNNGLPSNLSGGNPSLDYGFKGSEIAMASVSELOFLANPTN 480

QY 474 HVQASQHNODVNSLGLISARKTAFAVDILKIMSSYIYALQSIDLRHLENNKSTYKN 533
 DB 481 HVQASQHNODVNSLGLISARKTAFAVDILKIMSSYIYALQSIDLRHLENNKSTYKN 540

QY 534 TVSOVAKKVTGTVGNELHPSFCEKDLFRVYFVYFVYFVYFVYFVYFVYFVYFVY 593
 DB 541 TVSOVAKKVTGTVGNELHPSFCEKDLFRVYFVYFVYFVYFVYFVYFVYFVYFV 600

QY 594 DIALNNGETEKNTSIFOKIATFEELVLLPEKEVEGVRIAYENDTSLPPIKACRSY 653
 DB 601 DIALNNGETEKNTSIFOKIATFEELVLLPEKEVEGVRIAYENDTSLPPIKACRSY 660

QY 654 PLVRYVREHIGSGHFGKQKTSIGCEHFDVFTAMCKGQIDPLLEGLGQMGHPPIC 711
 DB 661 PLVRYVREHIGSGHFGKQKTSIGCEHFDVFTAMCKGQIDPLLEGLGQMGHPPIC 718

RESULT 11
 024266 PRELIMINARY: PRT, 715 AA.

AC 024266;
 DT 01-JAN-1998 (TREMblrel, 05, created)
 DT 01-JUN-1998 (TREMblrel, 05, last sequence update)
 DT 01-JUN-2002 (TREMblrel, 21, last annotation update)
 DE phenylalanine ammonia-lyase (EC 4.3.1.5).
 OS Populus kirkianaensis (Aspen).
 OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae, eucosids I; Malpighiales, Salicaceae; Populus.
 ON NCHI_taxid=1698;
 OX 11;
 RN SOURCE: FROM N.A.
 RP OSakabe Y., Ohnubo Y., Kawai S., Katayama Y., Morishashi N.;
 RT "Structures and tissue-specific expression of genes for phenylalanine
 RT ammonia-lyase from a hybrid aspen, Populus kirkianaensis.";
 RL Plant Sci. 105:217-226(1995).
 DB FMBI: D3056; MAA21643.1;

DR HSP: P21310; 188P;
 DR InterPro: IPR001106; Phc/His_NH3lyase.
 DR Pfam: PF00221; PAL, 1.
 DR TrEMBL: T16P01226; Phc-am_lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE, 1.
 KW lysase.
 SU SOURCE: 715 AA; 77626 MW; 040F098439H74DF C6C64;

Query Match 83.7%; Score 3049; DB 10; Length 715;
 Best Local Similarity 82.7%; Pred. No. 5,66,185;
 Matches 589, Conservative 58, Mismatches 61, Indels 4, Gaps 4;

QY 2 ENNRNNGVNNELTK-----DPLNMGVAEALTGSHDEVRKMAVEPRPVAKG 60
 DB 6 KNG-YONGSESCTQDRLPSKVAEAKKSHLDEVRKMAVEPRPVAKG 64

QY 61 VASIAANDSDTVKVELSEAPAPAVKASSDWMESNNKGTDSYVTTGKATSHPRTKQ 120
 DB 65 VASIAANDSDTVKVELSEAPAPAVKASSDWMESNNKGTDSYVTTGKATSHPRTKQ 123

QY 121 GADKELIRFLNAGIIGNCTETSHLPHSATPAAMIVRINTLQYSGIPEEI 180
 DB 124 GADKELIRFLNAGIIGNCTETSHLPHSATPAAMIVRINTLQYSGIPEEI 183

QY 181 LNNNTTCLPLRTTTASGDIYPLSTIAGLTGPPNSKAVPTGSEVLAKEAF 240
 DB 184 LNNNTTCLPLRTTTASGDIYPLSTIAGLTGPPNSKAVPTGSEVLAKEAF 243

QY 241 GFELOPKRELALVNIIVAGSGMASVLETAVALLESEVLSAIFAEVQKRP 300
 DB 244 GFELOPKRELALVNIIVAGSGMASVLETAVALLESEVLSAIFAEVQKRP 303

QY 301 HRTKNGGALQKELIFLNAGIFGNCTKSHLPHSATPAAMIVRINTLQYSGIPEEI 360
 DB 304 HRTKNGGALQKELIFLNAGIFGNCTKSHLPHSATPAAMIVRINTLQYSGIPEEI 363

QY 361 VIKSTKMEIREINSVNDPLIDVSNKALHGGNFQGTPIGVSMTKLATAA 420
 DB 364 VIKSTKMEIREINSVNDPLIDVSNKALHGGNFQGTPIGVSMTKLATAA 423

QY 421 QFSTLVNFTNNLPLNSLS:RNFSLTYGFKSEIAMASVSELOFLANPTN 480
 DB 424 QFSTLVNFTNNLPLNSLS:RNFSLTYGFKSEIAMASVSELOFLANPTN 483

QY 481 HNDVNSLGLISARKTAFAVDILKIMSSYIYALQSIDLRHLENNKSTYKN 533
 DB 484 HNDVNSLGLISARKTAFAVDILKIMSSYIYALQSIDLRHLENNKSTYKN 543

QY 540 KVLITMGVNGELHPSFCEKDLFRVYFVYFVYFVYFVYFVYFVYFVYFVY 599
 DB 544 KVLITMGVNGELHPSFCEKDLFRVYFVYFVYFVYFVYFVYFVYFVYFV 603

QY 600 GEFPEKNTSIFOKIATFEELVLLPEKEVEGVRIAYENDTSLPPIKACRSY 659
 DB 604 GEFPEKNTSIFOKIATFEELVLLPEKEVEGVRIAYENDTSLPPIKACRSY 663

QY 660 REELSRGELLIGKQKTSIGCEHFDVFTAMCKGQIDPLLEGLGQMGHPPIC 711
 DB 664 REELSRGELLIGKQKTSIGCEHFDVFTAMCKGQIDPLLEGLGQMGHPPIC 715

RESULT 12
 0944N1 PRELIMINARY: PRT, 725 AA.

AC 0944N1;
 DT 01-DEC-2001 (TREMblrel, 19, created)
 DT 01-JUN-2002 (TREMblrel, 21, last sequence update)
 DE Eukaryotic phenylalanine ammonia-lyase.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae;
 ON eucosids II; Brassicales, Brassicaceae, Arabidopsis.

```

OX NCB1_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Hower L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
RA Kawai J., Kim C., Kosuma E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Saitou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.P., Theologis A.,
PT Full length cDNA of gene T1J8.22/Atg37040 (c1:6598547).
PI Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
PM EMBL: AY045919; AAK76593.1;
DR InterPro: IPR001106; Phe/His_NH3lyase.
DR Pfam: PF00221; PAL: 1.
DR PRINTS: PR01568; LYMHOTACTNP
DR TIGRfams: TIGR01226; Phe_am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN_1.
KW lyase.
SQ
SEQUENCE 725 AA; 78725 MW; 026263B3DEFE9CE CRC64,
Query Match 83.6%; Score 3045.5; DB 10; Length 725;
Best Local Similarity 81.4%; Pred. No. 9.6e-185;
Matches 569; Conservative 52; Mismatches 66; Indels 17; Gaps 3;
OY 3 NGNHN---GVNNELC-----IKDPLNMVGAALGSHLDEVKKVAEERKP 47
DB 4 NGAHNSGGVADMLCGGDIKKNVINAEDPLNMGAAAEQKSHLDEVKVAEERKP 63
OY 48 VVKLGFTLTVSQVAGIAANDSTVKKVELEAARACAKSSIDVWMSMKCTDSCVTT 107
DB 64 VVNLGSETLTIGVAAISTGNS--VKVELSETARAGVNASDWMESMNGTDSYCVTT 121
OY 108 GFGATSHRRTKGSGALOKELIFELNAGIFGNATETSHLPISATRAAMIVKINTLLGVS 167
DB 122 GFGATSHRRTKGSGALOKELIFELNAGIFGNATETSHLPISATRAAMIVKINTLLGVS 181
OY 168 GTFEELTIAITFLNNITPCLPLRGITTAGSDLVLSYAGLLTGPENSKAVGPTGVL 227
DB 182 GTFEELTIAITFLNNITPCLPLRGITTAGSDLVLSYAGLLTGPENSKAVGPTGVL 241
OY 228 NAKKATGAAGVHGCFEELPKKGLAVNGTAVSGMASMVLFPDANVLAALLEVLSAIPAE 287
DB 242 TAEFAKKLGISSGFEDLPKGLAVNGTAVSGMASMVLFPDANVLAALLEVLSAIPAE 301
OY 288 VMGKPEFTDHLTHLKHHPQIEAAAMEYLLDGSDDVKAOKVHEMDLPQKPKODRYA 347
DB 302 VMGKPEFTDHLTHLKHHPQIEAAAMEYLLDGSDDVKAOKVHEMDLPQKPKODRYA 361
OY 348 LRTSPWMLGUVLEVRSSTKMIEKELNSVNDPLIDVSRKKAHGGNFQGTPIGVSMDNT 407
DB 362 LRTSPWMLGUVLEVRSSTKMIEKELNSVNDPLIDVSRKKAHGGNFQGTPIGVSMDNT 421
OY 408 RLALIAIGKLMFAOFSELVNDYNNGLPSNLGGKNSLDYGFKGGLTAMASVCSHIOPL 467
DB 422 RLALIAIGKLMFAOFSELVNDYNNGLPSNLGGKNSLDYGFKGGLTAMASVCSHIOPL 481
OY 468 ANEVTNHVOSADQHNODVNSLGLISARKTAEAVDILKMSSTLYVALCOSIDLRHLEENM 527
DB 482 ANEVTNHVOSADQHNODVNSLGLISARKTAEAVDILKMSSTLYVALCOSIDLRHLEENM 541
OY 528 KSTVKTQVAVAKKYLITKYNKELHPSRCECDLRYVDREYVAYIDDVCSGTPYLMOK 587
DB 542 RQVKTQVAVAKKYLITKYNKELHPSRCECDLRYVDREYVAYIDDVCSGTPYLMOK 601
OY 588 LKQVLDHALNGETEKNTNTSIFOKIATFEELKVLIPKVFVGRIAYENDTISTPNI 647
DB 602 LKQVLDHALNGETEKNTNTSIFOKIATFEELKVLIPKVFVGRIAYENDTISTPNI 661
OY 648 KACRSYPLRYFVREELISRGFTLTGEKVTSGGEFDFVFTAMCKGQJLDPLECTLGMMGEP 707
DB 662 KACRSYPLRYFVREELISRGFTLTGEKVTSGGEFDFVFTAMCKGQJLDPLECTLGMMGEP 721

```

```

OY 708 LPTC 711
DB 722 IFTC 725
RESULT 13
OBRWP4 PRELIMINARY: PRT: 717 AA.
AC OBRWP4: 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DE 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
GN Phenylalanine ammonia-lyase.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae;
OC eucosids 11; Brassicales; Brassicaceae; Arabidopsis.
OX NCB1_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamita A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Saitou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J.P., Theologis A., Davis R.W.,
PI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY02957; AAK12956.1;
KW lyase.
SQ
SEQUENCE 717 AA; 77801 MW; C4C9C8DCDB91A838 CRC64;
Query Match 82.6%; Score 3009; DB 10; Length 717;
Best Local Similarity 81.6%; Pred. No. 2e-182;
Matches 578; Conservative 52; Mismatches 76; Indels 2; Gaps 1;
OY 4 GNRHVGYNELCIKDPLNMVGAALGSHLDEVKKVAEERKPVYKLGSETLTVSQVAG 63
DB 12 GERTKVAATVTKTLADPLNMVGAALGSHLDEVKKVAEERKPVYKLGSETLTIGVAA 71
OY 64 IAAANDSTVKKVELEAARACAKSSIDVWMSMKCTDSCVTTGFGATSHRRTKKOGAL 123
DB 72 IFTVGS--VKVELSETARAGVNASDWMESMNGTDSYCVTTGFGATSHRRTKKOGAL 129
OY 124 OKELIFELNAGIFGNATETSHLPISATRAAMIVKINTLLGVSAGTFEELTIAITFLNN 183
DB 130 OKELIFELNAGIFGNATETSHLPISATRAAMIVKINTLLGVSAGTFEELTIAITFLNN 189
OY 184 NITPCLPLRGITTAGSDLVLSYAGLLTGPENSKAVGPTGVLNAKKAAGVHGCF 243
DB 190 NITPCLPLRGITTAGSDLVLSYAGLLTGPENSKAVGPTGVLNAKKAAGVHGCF 249
OY 244 ELQPKKGLAVNGTAVSGMASMVLFPDANVLAALLEVLSAIPAEVMOGKPEFTDHLTHL 303
DB 250 ELQPKKGLAVNGTAVSGMASMVLFPDANVLAALLEVLSAIPAEVMOGKPEFTDHLTHL 309
OY 304 KHHFGQIEAAAMEYLLDGSDDVKAOKVHEMDLPQKPKODRYALRTSPWMLGUVLEVR 363
DB 310 KHHFGQIEAAAMEYLLDGSDDVKAOKVHEMDLPQKPKODRYALRTSPWMLGUVLEVR 369
OY 364 SSTKMIEKELNSVNDPLIDVSRKKAHGGNFQGTPIGVSMDNTRLAIAIGKLMFAOFS 423
DB 370 SSTKMIEKELNSVNDPLIDVSRKKAHGGNFQGTPIGVSMDNTRLAIAIGKLMFAOFS 429
OY 424 ELVNDYNNGLPSNLGGKNSLDYGFKGGLTAMASVCSHIOPLANVTNHVOSADQHNOD 483
DB 430 ELVNDYNNGLPSNLGGKNSLDYGFKGGLTAMASVCSHIOPLANVTNHVOSADQHNOD 489
OY 484 DVNSLGLISARKTAEAVDILKMSSTLYVALCOSIDLRHLEENKSTVKTQVAVAKKYL 543
DB 490 DVNSLGLISARKTAEAVDILKMSSTLYVALCOSIDLRHLEENKSTVKTQVAVAKKYL 549

```



```

QY 356 GFOIEYIRSTKMEIEEINSVNDNPLIDVSRKKALHGGNFUGFPFGVSMONTRELATAIG 415
Db 300 GFOIEYIRHATMIREINSVNDNPLIDVSRKKALHGGNFUGFPFGVSMONTRELATAIG 359
QY 416 KLMFAGFSEIVNFFYNNMSPNLG3RNPSTLDYGFEGEIAMASYCSELOFLANPTNHV 475
Db 360 KLMFAGFSEIVNFFYNNMSPNLG3RNPSTLDYGFEGEIAMASYCSELOFLANPTNHV 419
QY 476 QSAPOHNOUVNLSGLISAFTFAVDILKMSSTYVALCOSIDLRIHEENKSTVANTV 535
Db 420 QSAPOHNOUVNLSGLISAFTFAVDILKMSSTYVALCOSIDLRIHEENKSTVANTV 479
QY 536 SOYAKKVLTMGVNGELHPSRCECKDLRVVDREYVFAYIDVCSGTYPLMOKLRQVLVDH 595
Db 480 SOYAKKVLTMGVNGELHPSRCECKDLRVVDREYVFAYIDVCSGTYPLMOKLRQVLVDH 539
QY 596 ALNNGETEKNTSIPUKIATFEELKVLLEKEVEGVRIAYENDTSLIPNRIKACRSYPL 655
Db 540 ALNNGETEKNTSIPUKIATFEELKVLLEKEVEGVRIAYENDTSLIPNRIKACRSYPL 599
QY 656 YRFVREELGRGFLTGKVTSPGEFDRVFTAMCKQIIDLPLHGLGGMNGEPLPIC 711
Db 600 YRFVREELGRGFLTGKVTSPGEFDRVFTAMCKQIIDLPLHGLGGMNGEPLPIC 655

```

Search completed: March 29, 2003, 02:12:14
 Job time : 67 secs